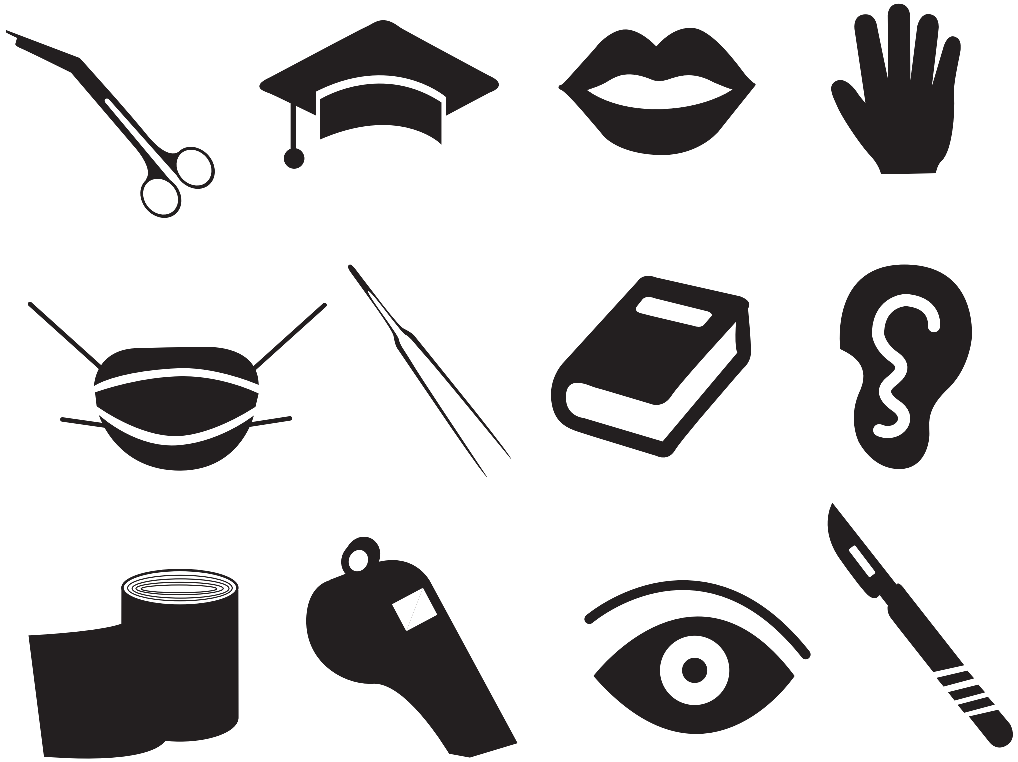


COACHING IN THE OPERATING ROOM

SANDER ALKEN

COACHING IN THE OPERATING ROOM

Sander Alken

The work presented in this thesis was carried out within the Radboud Institute for Health Sciences.

Financial support for the publication of this thesis was kindly provided by the Department of Surgery of the Radboud University Medical Center and the Radboud Institute for Health Sciences.

De uitgave van dit proefschrift is mede ondersteund door de Nederlandse Vereniging voor Medisch Onderwijs.

ISBN
978-94-028-1426-2

Design/lay-out
Bregje Jaspers, ProefschriftOntwerp.nl, Nijmegen

Print
Ipskamp printing

COACHING IN THE OPERATING ROOM

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 22 mei 2019
om 10.30 uur precies

door

Alexander Paulus Beatrix (Sander) Alken
geboren op 23 november 1984
te Geleen

Promotor
Prof. dr. H. van Goor

Copromotoren
Dr. C.R.M.G. Fluit
Dr. J.M. Luursema

Manuscriptcommissie
Prof. dr. M.A.W. Merkx
Prof. dr. J. de Graaf
Prof. dr. J.A. van der Hage (LUMC)

COACHING IN THE OPERATING ROOM

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Wednesday, May 22, 2019
at 10.30 hours

by

Alexander Paulus Beatrix (Sander) Alken
Born on November 23, 1984
in Geleen (The Netherlands)

Supervisor
Prof. dr. H. van Goor

Co-supervisors
Dr. C.R.M.G. Fluit
Dr. J.M. Luursema

Doctoral Thesis Committee
Prof. dr. M.A.W. Merkx
Prof. dr. J. de Graaf
Prof. dr. J.A. van der Hage (Leiden University Medical Center)

TABLE OF CONTENTS

CHAPTER 1

General introduction

Technical and non-technical skills training

Surgical training has always had a strong focus on developing high quality technical skills.[1] Technical skills in surgery encompass psychomotor skills like handling a scalpel and forceps, suturing, dissecting or resecting organs and other structures, and implanting prostheses. Although technical skills are highly important, studies show that most errors and mistakes made in surgery are not caused by deficiencies in surgeons' technical skills, but rather in non-technical team and communication skills.[2,3] Non-technical team and communication skills encompass, for example, the assisting of scrub nurses in preparing the correct instruments or the sharing of relevant information with the anesthesiologist regarding operative complications. High quality and safe surgery requires the concurrent application of effective technical and non-technical skills.

Although the attention for non-technical skills is rising, it is only a minimal, infrequent and unstructured part of surgical training.[3,4] Teaching of non-technical skills in the operating room is coincidental, mostly when such situations occur.[4] Furthermore, surgeons are unacquainted with the teaching of non-technical skills and naturally focus on technical skills,[3] possibly due to lack of specific teaching skills for non-technical skills. As a result, non-technical skills remain under taught. To train surgeons effectively in technical and non-technical skills and prepare them for independent practice, it has been suggested that non-technical skills should be taught and trained next to technical skills[5-8] and preferably concurrently.[9] How to effectively combine non-technical with technical skills teaching is largely unknown.

Teaching behavior

What behaviors surgical teachers apply when they are coaching trainees seems to have important consequences for the effectiveness of (non-)technical skill training. Providing instructions, feedback and explanations, for example, are considered effective coaching behaviors.[10] However, if provided too intensively, coaching may become ineffective because it may cause information and cognitive overload in trainees. Overload distracts trainees from learning and practicing the skill[10,11] and makes trainees dependent on coaching which hampers the development to independent practice.[12] Coaching behaviors that teachers apply in the operating room can be labeled as demonstrations, instructions, feedback, explanations, questions and conversations.[13-17] There is little insight in the specific content of coaching provided by the teachers and the balance between coaching on technical and non-technical skills.

The operating room as teaching environment

Surgical trainees train surgical skills mostly in the real-life operating room while operating on patients under the supervision of experienced surgeons. In this high-stakes, complex and dynamic environment trainees should master all skills necessary to become a surgeon. From an educational perspective, however, the operating room is not an ideal teaching environment. The patient's safety is always the number one priority which limits the possibilities and opportunities for teaching.[18] Surgical procedures can be complex, complications may occur and the pressure to finish the operation in time is often high.[19] These elements can easily cause stress in surgeons[20] which is known to negatively affect communication styles[19,21-24] and evoke disruptive communication.[21] If surgical teachers apply disruptive communication styles it jeopardizes the learning of surgical trainees.[25] Trainees often report to be unsatisfied with the quality of coaching.[26-30] Surprisingly, the relation between stress in teachers and their coaching behavior has not yet been topic of surgical educational research.

The time available for teaching in the operating room is increasingly diminishing due to demands for efficient patient care, more complex operations and surgical techniques, and, in some countries, significant limitations in trainees' working hours.[20,30,31] To date, trainees feel insufficiently prepared for independent practice.[30] These limitations, together with the complex and hectic circumstances in the operating room, urge the need for highly skilled and effective surgical teachers in both technical and non-technical skills. Increasing our understanding of how surgical teachers coach during complex procedures and how stress relates to their coaching behaviors may contribute to the development of new strategies for the effective coaching of surgical trainees.

Simulation as teaching environment

Surgical skills can also be taught and trained outside the operating room through simulation. The models used in surgical simulation training vary and encompass, amongst others, bench top models, virtual reality models, human manikin models, animal cadaver models, human cadaver models or live animal models.[32,33] Each model has advantages regarding training time, teaching strategies that can be applied, training of technical and non-technical skills, possibilities to freeze and pause situations, and the reality of the training environment. Bench top models, for example, offer trainees and teachers to train (simple) technical skills in an exploratory way without time limits or pressure. If live animal models are involved, simulation becomes highly realistic since injuries, bleedings and complications are real and cannot be paused.[32] Such a simulation environment is a hectic teaching environment which closely resembles the operating room as a teaching environment.

Research questions

In our studies we investigated what teaching behaviors teachers apply during a live animal simulation team training which closely resembles the operating room as a teaching environment. The pivotal questions which we attempt to answer in this thesis are:

1. What behaviors do surgical teachers apply to teach surgical trainees in the operating room? More specifically:
    a. How do surgical teachers apply feedback, instructions and detailed explanations when teaching technical and non-technical skills during a highly realistic simulated team training?
    b. How do surgical teachers and trainees perceive the applied feedback, instructions, detailed explanations and coaching on technical and non-technical skills during a highly realistic simulated team training? And how do these perceptions relate to independent observations?
    c. What are evidence-based coaching behaviors to effectively teach complex technical psychomotor skills during hands-on training situations?

2. How can teaching of non-technical skills be effectively combined with teaching of technical skills in the operating room? More specifically:
    a. Can priming of surgical teachers increase the frequency of coaching on non-technical skills during a highly realistic simulated team training?

3. What similarities or differences in stress patterns do surgeons and residents display during different activities in daily surgical practice?

4. How does stress in surgical teachers relate to the teaching behaviors they apply in the operating room?
    a. Do surgical teachers apply feedback, instructions, detailed explanations and coaching on technical and non-technical skills during moments of highest stress differently from moments of lowest stress, during a highly realistic simulated team training?
    b. How does stress in a surgical teacher and a surgical trainee relate to specific events occurring in the operating room during a highly realistic simulated team training?

Research context and methods

Six studies were conducted at the Department of Surgery of the Radboud University Medical Center in Nijmegen, the Netherlands. The Department of Surgery provides training for medical students (bachelor and master program), residents (postgraduate training program) and consultant surgeons (advanced training program). An advanced training course, the annually held Definitive Surgical and Anesthetic Trauma Care (DSATC) course, was the primary setting for most studies.

The Definitive Surgical and Anesthetic Trauma Care course

The DSATC course is a three-day, highly realistic acute trauma surgery team training, combining the Definitive Surgical (DSTC) and Definitive Anesthetic Trauma Care (DATC) course.[34,35] For surgical trainees, the primary objectives of the DSATC course are the training of surgical technical skills (damage control surgery techniques) and multidisciplinary team and communication skills necessary to provide effective care for critically injured patients. The course is annually held and accessible for final-year surgical residents, fellows and consultant surgeons with an interest in trauma surgery.

Each of the three days starts with plenary lectures in the morning and practical workshops in the afternoon. All lectures focus on technical surgical skills and surgical decision making, except for one 30-minute lecture focusing on non-technical team and communication skills. The lectures are attended by all teachers and trainees together. Three parallel practical workshops, each lasting the entire afternoon, are held every day. In three days each trainee completes all three practical workshops. One workshop is the damage control surgery team training in the Central Animal Laboratory.

During this highly realistic simulated team training, one certified and experienced surgical teacher coaches two surgical trainees while operating in an operating room setting with an anesthetic team and a scrub nurse. The anesthetic team is composed of an experienced anesthesiologist, who is also a certified teacher, and who teaches emergency anesthesia to one or two trainees during the training. The research described in this thesis focuses on the coaching behaviors of the surgical teachers. During this workshop the teacher inflicts six consecutive complex and life-threatening injuries on anesthetized porcine models. Surgical trainees are unaware of the type and location of each injury and have to gain control of the "patient" in each of the six trauma scenarios. The DSATC training is approved by the animal ethical review board. For each study the teachers and trainees gave permission for obtaining and using their data.

General description of research methods

Several quantitative research methods and one qualitative research method are applied in this thesis. All activities during the damage control surgical team training workshops were recorded on video by cameras placed on a tripod in the operating room (Figure 1) and on audio by voice-recorders placed above the operating area or attached to the gowns of the surgical teachers and/or trainees at the shoulder height. To analyze the coaching behaviors of surgical teachers we designed a qualitative observational instrument based on literature reports and expert input.

Figure 1. Screenshot made with a video camera during the damage control surgery team training.

Thesis objectives

The objectives of this thesis are:

1. to categorize and analyze the coaching behaviors surgical teachers apply when teaching in the operating room;
2. to determine how these teaching behaviors are perceived by surgical teachers and trainees, and how they relate to independent observations;
3. to investigate if the priming of teachers results in an increase in non-technical skills coaching;
4. to determine stress of surgeons and residents in daily surgical practice;
5. to determine how surgical teachers experience stress during teaching in the operating room and how this relates to their coaching behaviors;
6. to review evidence-based coaching behaviors that are effective in teaching complex psychomotor skills;
7. to provide an overview on how to integrate technical next to non-technical skills teaching and training in the operating room.

Thesis outline

Chapter 1 contains the general introduction. In **Chapter 2** we analyze the teaching behaviors that surgical teachers apply in the operating room during the simulated damage control surgery team training. We compare the frequency of instructions to the frequency of feedback, the frequency of technical skills coaching to non-technical skills coaching and compare different levels of specific information provided by the teachers. In **Chapter 3** the perceptions of teachers and trainees are compared regarding the provided teaching and the observations that were recorded during the team training. In **Chapter 4** the effect on coaching behavior of an additional instruction session with the purpose to prime teachers on non-technical skills' coaching, is assessed and compared with behavior of teachers who had no priming session. In **Chapter 5** objective and subjective stress of consultant surgeons and surgical residents is investigated during daily activities and risk factors for increased stress are determined. The objective stress is measured by heart rate variability (HRV) using an innovative patch. In **Chapter 6** we describe and analyze patterns of teaching in the trauma course during periods of highest and lowest stress using the same patch as in the study described in chapter 5. **Chapter 7** contains a systematic review of existing evidence on effective teaching of technical skills in surgery and in other fields of psychomotor skills teaching. In **Chapter 8** we summarize the state of the art knowledge regarding the hands-on teaching of technical and non-technical skills in surgery, which we partly base on the studies we conducted. We also propose a theoretical framework focusing on how to integrate the teaching of both skill sets in modern surgical training. We close with expressing our thoughts for future research concerning effective hands-on surgical coaching. **Chapter 9** contains the summaries of the studies in this thesis.

REFERENCES

1. Agha RA, Fowler AJ and Sevdalis N, The role of non-technical skills in surgery. Annals of Medicine and Surgery, 2015. 4(4): p. 422-427.
2. Gjeraa K, et al., Non-technical skills in minimally invasive surgery teams: a systematic review. Surgical endoscopy, 2016. 30(12): p. 5185-5199.
3. Hull L and Sevdalis N, Advances in Teaching and Assessing Nontechnical Skills. Surg Clin North Am, 2015. 95(4): p. 869-84.
4. Spanager L, et al., Comprehensive feedback on trainee surgeons' non-technical skills. International journal of medical education, 2015. 6: p. 4.
5. Brunckhorst O, et al., The relationship between technical and nontechnical skills within a simulation-based ureteroscopy training environment. Journal of surgical education, 2015. 72(5): p. 1039-1044.
6. Dawe S, et al., Systematic review of skills transfer after surgical simulation-based training. British Journal of Surgery, 2014. 101(9): p. 1063-1076.
7. Dedy NJ, et al., Teaching nontechnical skills in surgical residency: a systematic review of current approaches and outcomes. Surgery, 2013. 154(5): p. 1000-1008.
8. Whittaker G, et al., Teamwork assessment tools in modern surgical practice: a systematic review. Surgery research and practice, 2015. 2015.
9. Azuara-Blanco A, et al., Safe eye surgery: non-technical aspects. Eye (Lond), 2011. 25(9): p. 1109-11.
10. White C, Rodger MW and Tang T, Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. The Obstetrician & Gynaecologist, 2016. 18(1): p. 53-63.
11. Nicholls D, et al., Teaching psychomotor skills in the twenty-first century: revisiting and reviewing instructional approaches through the lens of contemporary literature. Medical teacher, 2016. 38(10): p. 1056-1063.
12. Walsh CM, et al., Concurrent versus terminal feedback: it may be better to wait. Acad Med, 2009. 84(10 Suppl): p. S54-7.
13. Blom EM, et al., Analysis of verbal communication during teaching in the operating room and the potentials for surgical training. Surg Endosc, 2007. 21(9): p. 1560-6.
14. Chen XP, et al., A taxonomy of surgeons' guiding behaviors in the operating room. Am J Surg, 2015. 209(1): p. 15-20.
15. Hauge LS, Wanzek JA, and Godellas C, The reliability of an instrument for identifying and quantifying surgeons' teaching in the operating room. Am J Surg, 2001. 181(4): p. 333-7.
16. Roberts NK, et al., Capturing the teachable moment: a grounded theory study of verbal teaching interactions in the operating room. Surgery, 2012. 151(5): p. 643-50.
17. Sutkin G, Littleton EB, and Kanter SL, How surgical mentors teach: a classification of in vivo teaching behaviors part 1: verbal teaching guidance. J Surg Educ, 2015. 72(2): p. 243-50.
18. Kneebone R, et al., Distributed simulation--accessible immersive training. Med Teach, 2010. 32(1): p. 65-70.
19. Arora S, et al., Factors compromising safety in surgery: stressful events in the operating room. The American Journal of Surgery, 2010. 199(1): p. 60-65.

20. Goldberg MB, et al., Optimizing performance through stress training - An educational strategy for surgical residents. Am J Surg, 2017.
21. Arora S, et al., Managing intraoperative stress: what do surgeons want from a crisis training program? The American Journal of Surgery, 2009. 197(4): p. 537-543.
22. Arora S, et al., The impact of stress on surgical performance: a systematic review of the literature. Surgery, 2010. 147(3): p. 318-330. e6.
23. Hull L, et al., The impact of nontechnical skills on technical performance in surgery: a systematic review. Journal of the American College of Surgeons, 2012. 214(2): p. 214-230.
24. Wetzel CM, et al., The effects of stress on surgical performance. The American Journal of Surgery, 2006. 191(1): p. 5-10.
25. Flinn JT, et al., The effect of stress on learning in surgical skill acquisition. Medical teacher, 2016. 38(9): p. 897-903.
26. Butvidas LD, et al., Disparities between resident and attending surgeon perceptions of intraoperative teaching. The American Journal of Surgery, 2011. 201(3): p. 385-389.
27. Chen X, Williams RG, and Smink DS, Do Residents Receive the Same OR Guidance as Surgeons Report? Difference Between Residents' and Surgeons' Perceptions of OR Guidance. Journal of Surgical Education, 2014. 71(6): p. e79-e82.
28. Levinson KL, et al., Disparity Between Resident and Attending Physician Perceptions of Intraoperative Supervision and Education. Journal of Graduate Medical Education, 2010. 2(1): p. 31-36.
29. Rose JS, Waibel BH, and Schenarts PJ, Disparity Between Resident and Faculty Surgeons' Perceptions of Preoperative Preparation, Intraoperative Teaching, and Postoperative Feedback. Journal of Surgical Education, 2011. 68(6): p. 459-464.
30. Anderson CI, et al., Impact of objectively assessing surgeons' teaching on effective perioperative instructional behaviors. JAMA Surg, 2013. 148(10): p. 915-22.
31. DaRosa DA, et al., A Theory-Based Model for Teaching and Assessing Residents in the Operating Room. Journal of Surgical Education, 2013. 70(1): p. 24-30.
32. Palter VN and Grantcharov TP, Simulation in surgical education. Canadian Medical Association Journal, 2010. 182(11): p. 1191-1196.
33. Sadideen H, et al., Simulators and the simulation environment: Getting the balance right in simulation-based surgical education. International Journal of Surgery, 2012. 10(9): p. 458-462.
34. Bleeker CP, Definitive Anaesthetic Trauma Care Courses. [cited 2016 2016-11-7]; Available from: http://www.iatsic.org/DATC.html.
35. Unknown. Definitive Surgical and Anaesthetic Trauma Care Course [cited 2016 2016-11-7]; Available from: http://www.dstc-datc.nl/media/192349/folder_dsatc-datc_2016.pdf.

CHAPTER 2

Feedback activities of instructors during a trauma surgery course

Alexander Alken[1], Edward Tan[1], Jan-Maarten Luursema[1], Cornelia Fluit[2], Harry van Goor[1]

[1]Department of Surgery
[2]Radboudumc Health Academy

Radboud University Medical Center, Nijmegen, The Netherlands

The American Journal of Surgery 2013; 206 (4): 599-604

ABSTRACT

BACKGROUND: The aim of this study was to examine the quality and quantity of feedback and instruction from faculty members during an acute trauma surgery team training using a newly designed observational feedback instrument.

METHODS: During the training, 11 operating teams, each consisting of 1 instructor coaching 2 trainees, were videotaped and audio taped. Forty-five minutes of identical operating scenarios were reviewed and analyzed. Using a new observational feedback instrument, feedback and instruction, containing different levels of specific information related to technical and non-technical skills, were noted.

RESULTS: Instructors more often provided instruction (25.8 ± 10.6 times) than feedback (4.4 ± 3.5 times). Most feedback and instruction contained either nonspecific or less specific information and referred to technical skills. Instructors addressed communication skills more specifically.

CONCLUSIONS: Coaching by faculty members predominantly contained unspecific instructions regarding technical skills. The observational feedback instrument enabled scoring of the coaching activities.

INTRODUCTION

Instructor feedback is essential for acquiring surgical skills in practical training situations.[1–4] To be effective, instructors should provide trainees with well-timed feedback that contains specific information to improve the trainees' performance, knowledge, and skills.[1,4,5]

Specific feedback includes information on how trainees are doing, what went well, what needs to be improved,[2,5] how improvements can be accomplished, and which alternative approaches exist.[5] The most effective feedback informs trainees whether a performance was correct or incorrect, enriched with specific information on how trainees should have performed and supported by an explanation why this is so.[6,7]

The timing for feedback is either during training, while trainees are still performing (immediate feedback), or after training (delayed feedback).[8] Teaching during surgery is an important aspect of surgical training and generally includes immediate feedback.[9,10] Immediate feedback enables instructors to provide instant correction and encouragement during the training procedure.[10,11] It has the advantage of training the correct skills and behavior immediately and prevents trainees from automating incorrect skills and behavior.[11] When applying immediate feedback, however, instructors risk using step-by-step and unspecified instruction instead of feedback, with possible cognitive overload of trainees.[8,12]

The majority of training courses in surgery are more focused on the technical and less on the non-technical communication and team skills.[13–16] This is understandable for simple training such as suturing but not for training complex surgical skills that involve crew resource management. Notably, many errors in operating rooms are due to failures in non-technical skills.[13,14,17–19]

The Definitive Surgical Trauma Care (DSTC) course contains complex technical and non-technical skill exercises for a surgical and anesthesia team on anesthetized porcine models.[20] Life-threatening scenarios are simulated wherein traumatic injuries need to be controlled quickly and adequately. The team is subjected to stress and time constraints and may not make gross errors. To date, the feedback given in such a complex training situation has not yet been analyzed. The present study was designed to gain insight into the feedback given by instructors of the DSTC course with respect to the use of instruction, different levels of specific information, and technical and non-technical skills. The purpose of our study was to describe how instructors guide their trainees during the practical training of complex surgical skills.

METHODS

Participants

Eleven surgical instructors participated in the study. They were all experienced trauma surgeons and certified advanced trauma life support instructors with 2 to 12 years of teaching experience. Twenty-two Dutch surgical registrars participated as trainees. During the exercises, each instructor provided feedback to 2 surgical trainees who, with an operating room nurse, formed an operating team. An anesthesiologist trainee, instructed by an anesthetist, completed the team but was not included in this study. All operating teams were composed on the basis of alphabetic order of surname. Before the start of the training course, written informed consent from the instructors and the trainees participating in the feedback study was obtained. The institutional review board of Radboud University Nijmegen Medical Center waived the need for formal ethical approval for this educational study.

Surgical skill exercises

The study was performed during the emergency surgical trauma skill exercises on anesthetized pigs that are part of the annual 3-day DSTC course in the Netherlands. The DSTC combines interactive theoretical sessions with practical training on live porcine models and human corpses.

The exercises took place in 4 operating rooms in the central animal facility at our university. The trainees had to control complex and multiple abdominal and thoracic injuries in 6 acute emergency scenarios within 3.5 hours. Trainees were blinded to the type of injury that had been inflicted. The local animal ethical review board approved this training.

Study design

During all 6 scenarios, real-time audio and video recordings were made of each surgical team's operating activities. Of the 210 minutes recorded, 45 minutes were selected containing the same 2 scenarios for each team: (1) a stab wound to the infra-renal caval vein in the right mid-abdomen; and (2) a pericardial stab wound and injury to the left lung. In these scenarios, the technical and non-technical skills were considered to be important.

A trained observer analyzed the 45-minute recordings to register the feedback content using an observational instrument (Table 1). This checklist for surgical skills training observation was compiled on the basis of literature research and focused on instructions and reflective feedback,[8,9,11,12,21] specificity levels of feedback,[2,5,7] and feedback on technical and non-technical skills.[13–16] The checklist's main categorization was between feedback and instruction. Feedback had to be more reflective, to refer to previous executed actions, and to contain information about the (in)correctness of the performance,[2,8,12,21] such as ''Well done!'' Instruction had a steering character and directed the trainee what to do,[9] for example, ''First, check the entire organ.''

The main categories of feedback and instruction were divided into 4 subcategories within 3 domains. The subcategories distinguished between the specific, incremental levels of information.

According to previous research, information regarding skills can be directed at how a certain skill or action needs to be performed and why it has to be performed (in a certain way) or why other approaches are better.[6,7] On the basis of this information, we designed the following incremental levels of information specificity concerning feedback and instructions: not at all reinforced (considered least specific), reinforced with a ''how'' statement (moderately specific), reinforced with a ''why'' explanation (highly specific), and reinforced with both a ''how'' statement and a ''why'' explanation (extremely specific) (Table 1). The 3 domains were technical skills (suturing skills, instrument handling, etc), non-technical communication skills (informing the anesthesiologist, discussing possible approaches with team members, etc), and team cooperation skills (helping with handing instruments, assisting others, etc). Non-technical skills relate to abilities such as situational awareness, decision making, communication and teamwork, and leadership.[13,16] Team-related skills refer to communication skills such as sharing information, communicating options and actions to other team members, and team cooperation skills such as managing team activities, supporting others, and discussing options with other team members.

An example of the subcategory ''not at all reinforced'' referring to feedback on technical skills would be ''Terrible suturing!'' Other than telling the trainee that his or her suturing is incorrect, it contains no other information as to why the suturing is terrible and how it can be improved. An example in the subcategory ''reinforced with a 'how' statement'' pertaining to instructions on communication skills would be ''Never start a procedure without asking the anesthesiologist to update on the condition of the patient.'' It teaches the trainee how to communicate with the anesthesiologist but does not provide a ''why'' explanation. An example in the subcategory ''reinforced with a 'why' explanation'' pertaining to instructions on team cooperation skills would be ''Help each other. Helping your colleague saves time.'' It informs the trainee why helping a colleague is a good decision but not exactly how to help his colleague. Feedback in the subcategory ''reinforced with both a 'how' statement and a 'why' explanation'' would be ''Your suturing needs to be better. Make nice, small stitches closer to each other so that the wound will heal better.'' It tells the trainee to suture correctly and explains why it is important. This subcategory contains the most specific information.

Data analysis

Per instructor, each occurrence of feedback or instruction was scored. Feedback by the instructors was limited. Therefore, feedback and instructions were taken together to enable further analysis of level of specificity (see the 4 subcategories) and type of skill (see the 3 domains). Unintelligible feedback or instruction because of interfering noise was discarded, as were remarks irrelevant to the training. When 1 sentence contained feedback as well as instruction, contained different specificity levels, or addressed different skills, each part was scored independently. We focused on verbal feedback and did not register the duration of time spent on feedback, instruction, the different specificity levels, and the different skills. We also did not score nonverbal activities, assuming that trainees were focused on the operation and might not notice these from instructors.

The total mean scores for the 11 instructors were calculated on the basis of the total instructor scores on feedback and instructions, the 4 levels of specificity, and 3 types of skills. In this way, it was possible to determine the average frequency with which each instructor gave feedback and instructions to different levels of specificity for each type of skill. Finally, the mean differences between the total mean scores were calculated.

Table 1. The instrument for surgical skills training observation

	FEEDBACK[a]				INSTRUCTIONS[b]			
	Not reinforced at all	Reinforced with "how" statement	Reinforced with "why" explanation	Reinforced with "how" statement and "why" explanation	Not reinforced at all	Reinforced with "how" statement	Reinforced with "why" explanation	Reinforced with "how" statement and "why" explanation
Technical skills								
Communication skills								
Team cooperation skills								

[a] Information about the performance of the trainee during practical training (often accompanied by a value judgment). [b] Information on what to do next that instructs the trainee during practical training.

Statistical analysis

Descriptive statistics analysis, paired-samples Student's t tests for normally distributed variables, and nonparametric Wilcoxon's signed-rank tests for variables not normally distributed were used to analyze differences in the total amount of feedback and instructions, differences in the use of different specificity levels of reinforcement, and differences in the skills addressed by the 11 instructors. Correlations were tested between the frequencies in the different specificity levels and the types of skills, using Pearson's r for parametric and

Spearman's r for nonparametric samples. We used SPSS version 18.0 (SPSS, Inc, Chicago, IL) for all statistical analyses. *P* values < .05 were considered significant.

RESULTS

All relevant feedback activities of the instructors could be classified using the instrument for surgical skills training observation. Most comments made by the instructors could be scored; approximately 20% could not be scored because they were irrelevant, not related to the training, or inaudible. These were discarded.

Table 2 shows the descriptive statistics of the 45-minutes video and audio reviews for each of the 11 instructors. Instructors provided instruction (25.8 ± 10.6 times) significantly more often than feedback (4.4 ± 3.5 times) (t = -8.77, P = .000). Instruction was correlated positively with feedback (r = .79, P = .004): instructors who gave more instruction also gave more feedback.

Most frequent were unreinforced feedback and instruction (11.7 ± 5.2 times) and those reinforced with ''how'' statements (12.5 ± 7.0 times). Less frequent were those reinforced with ''why'' explanation (1.7 ± 2.0 times) or reinforced with both components combined (4.3 ± 5.7 times) (Table 2). The frequency of reinforcements with ''why'' explanations was not significantly greater than that of statements with both components combined (P = .14). Feedback and instruction reinforced with ''how'' statements were correlated positively with reinforcements containing both ''how'' and ''why'' explanations (ρ = .75, P = .01): instructors who gave more ''how'' statements also gave more reinforcements combining both components.

Instructors gave more feedback and instruction pertaining to technical skills (28.2 ± 12.3 times) than communication skills (1.7 ± 1.7 times) (t = 7.66, P = .000) and also fewer team cooperation skills (.3 ± .6 times) (P = .003, Wilcoxon's signed-rank test). Feedback and instruction on communication skills were more frequent than on team skills (P = .02, Wilcoxon's signed-rank test).

Feedback and instruction on communication skills were correlated positively with the more specific feedback and instructions: reinforced with a ''why'' explanation (r = .62, P = .04) and reinforced with both components combined (ρ = .60, P = .05).

COMMENTS

The faculty members of a complex acute trauma surgical skills training used unspecified instruction rather than reflective feedback, referring mostly to technical skills. There seemed to be patterns for instruction and feedback with poor and rich content among the instructors; some instructors gave more specific information than others (Table 1). With the observational instrument, it was possible to classify all relevant verbal coaching activities.

Table 2. Instructors' feedback and instruction divided into 4 subcategories (level of specificity of information) and within the 3 domains (skills)

	Category		Subcategory				Domain		
Instructor	*Feedback*	*Instruction*	*Not at all*	*How*	*Why*	*How and why*	*Technical skills*	*Communication skills*	*Team cooperation skills*
1	9	39	7	18	4	19	43	3	2
2	4	22	16	8	0	2	23	3	0
3	6	36	13	18	6	5	37	5	0
4	11	33	10	26	0	8	44	0	0
5	0	12	4	7	1	0	12	0	0
6	3	17	12	6	1	1	19	1	0
7	6	36	22	11	2	7	39	3	0
8	1	13	6	5	3	0	13	0	1
9	1	13	9	5	0	0	14	0	0
10	2	33	15	14	2	4	33	2	0
11	5	30	15	19	0	1	33	2	0
Mean ± SD	4.4 ± 3.5	25.8 ± 10.6	11.7 ± 5.2*	12.5 ± 7.0†	1.7 ± 2.0*,†	4.3 ± 5.7*,†	28.2 ± 12.3	1.7 ± 1.7	0.3 ± 0.6

* t = 5.79, mean difference = 10.0, P = .000, SD = 5.7: feedback and instruction not reinforced at all (mean, 11.7) compared with reinforced with a ''why'' explanation (mean, 1.7); mean difference = 7.4, P = .03, Wilcoxon's signed-rank test, SD = 7.7: feedback and instruction not reinforced at all (mean, 11.7) compared with reinforced with both a ''how'' and a ''why'' explanation (mean, 4.3).

† t = 5.08, mean difference = 10.8, P = .000, SD = 7.0: feedback and instruction reinforced with a ''why'' explanation (mean, 1.7) compared with those with ''how'' statements (mean, 12.5); mean difference = 8.2, P = .004, Wilcoxon's signed-rank test, SD = 6.0: feedback and instruction reinforced with a ''how'' statement (mean, 12.5) compared with both a ''how'' and a ''why'' explanation (mean, 4.3).

Evidence-based information concerning the appropriate combination of instruction and feedback is lacking for surgical simulation training. We were surprised to find that instructions to trainees constituted the main part of the instructors' feedback activities during this advanced course. Strict guidance is expected when novices are practicing surgical skills[9] or when instructors themselves are relatively inexperienced at providing feedback. Although the trainees in this course were experienced surgeons, they have had limited exposure to multiple injury patients, in particular those with thoracic and abdominal stab wounds. Such relative inexperience may have elicited instruction rather than feedback in a life-threatening situation with the risk of losing the animal. All instructors were experienced faculty members and at a minimum had successfully attended the advanced trauma life support ''train the trainer'' course, which, however, is an acute trauma life support course lacking the complex surgical elements of the DSTC. The apparent necessity felt to provide immediate feedback could have tempted faculty members to use step-by-step instructions to achieve good performance.[8,12] Step-by-step instructions are a drawback of immediate feedback, known as the guidance hypothesis. That is, the use of immediate feedback could lead to the instructors' guiding and instructing the trainees step by step to correct their performance, a process that decreases learning outcomes.[8,12] Such feedback is not reflective but of an instructive nature. We did not formally inventory the learning and teaching expectations of the trainees or their instructors regarding guidance intensity before starting the exercises. Clarification of these expectations might have improved the balance between instructions and feedback.

Most often, feedback and instruction were either not reinforced at all or were reinforced only with ''how'' statements, both indicating low specificity. It is known that time constraints, present in this training, do lower the specificity of feedback.[4] Instructors may also have refrained from highly specific feedback or instructions judging that participants were technically skilled for this training. For high-level trainees, feedback does not necessarily have to be specific to be effective: a ''good'' or a ''wrong'' could provide enough feedback information.[12,21] If instructors estimated the participants as being advanced and skilled, the question remains why instruction, not feedback, did dominate the coaching.

Instructors rarely addressed non-technical communication and team cooperation skills, although these skills were objectives of the course. We do not believe that the choice not to emphasize teaching non-technical skills was because participants were experienced in this area. Time constraints to repair the life-threatening injuries to the animals during the exercises seem to have created a sense of urgency; this better explains the focus on technical skills. Possibly it was simply too great a burden for instructors to simultaneously teach non-technical and complicated technical skills.[8,22,23] Alternatively, low awareness of the importance of these skills by the faculty member might account for this result. Individual scores and significant correlations suggest that some instructors were more focused on non-technical skill teaching and on giving specific feedback than others. One strategy to increase an instructor's awareness of the importance of non-technical skills during practical training sessions is cross-training.[24]

During cross-training, members of operating teams are trained to perform one another's tasks and responsibilities. When a surgical instructor is trained for the role of anesthesiologist, he experiences the wants and the needs and the desire for communication and teamwork of anesthesiologists during surgery. Timeout and stop-and-check pauses during such a surgical trauma care training may also enhance feedback on communication and team skills.[25] Another method to increase both the quantity and quality of feedback on technical and non-technical skills is verbalization by the trainees performing the exercise.[5]

We currently plan to study stop-and-check pauses in the DSTC course, in which delayed feedback on non-technical communication and team skills will be given per training scenario executed. Thus, the feedback approach will be immediate feedback on technical skills during the scenarios and delayed feedback on non-technical skills after each scenario. That is because, although immediate feedback enables faculty members to correct and encourage trainees instantaneously[10,11] and prevent them from automating skills incorrectly,[11] other researchers posit that immediate feedback may impair learning activities and skills acquisition.[8,12,21] Using immediate feedback involves the risk for cognitive overload of trainees when learning new skills and simultaneously receiving as well as understanding and responding to the feedback on their performance.[8,22,23] Instructors may also experience cognitive overload when assessing trainees' performance and generating immediate feedback while teaching new skills at the same time.[8] Overload particularly jeopardizes the specificity of the feedback that is aggravated by time constraints.[4]

Although we noticed differences in behaviors and teaching styles among instructors, we choose to sum the data from all instructors to provide significant numbers. Knowledge of different teaching styles and how teaching styles relate to learning styles of trainees is important when studying the learning effect of the training. Matching teachers and trainees on the basis of their learning styles seems to improve outcomes.[26–28] In future research, we will use learning style questionnaires to identify the preferences of the individual trainees and trainers for better goal directed teaching and learning.[29]

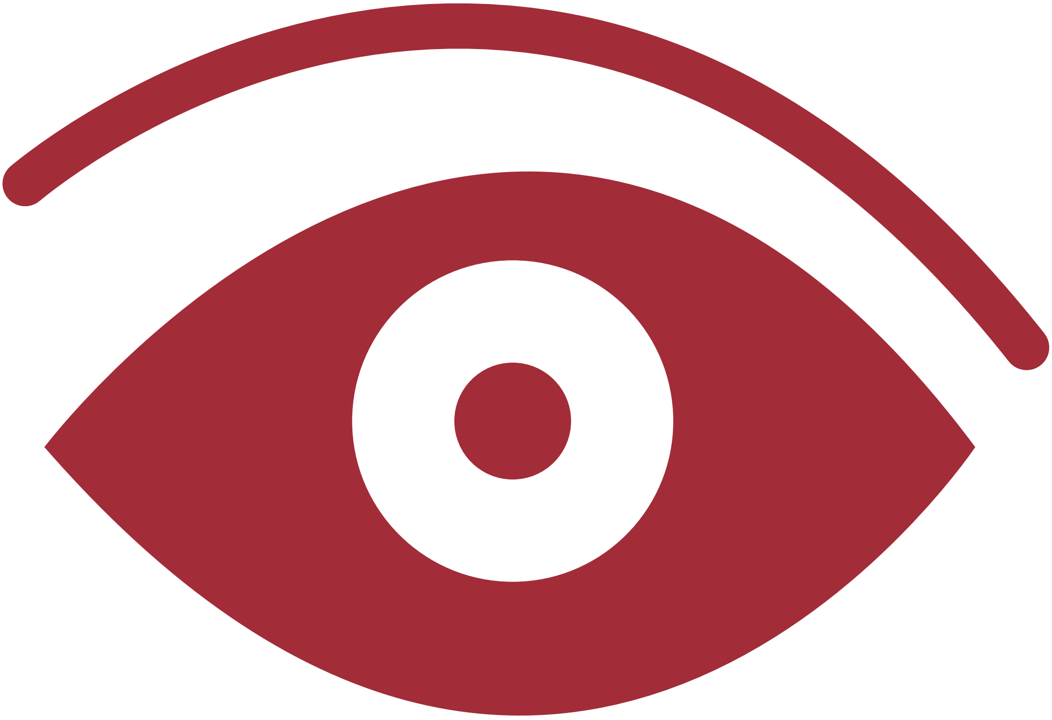
This study had some limitations. First, only one-third of scenarios were reviewed and analyzed. Although we felt the chosen scenarios to be representative, the order of scenarios might have influenced the amount and specificity of feedback and instruction. The observational instrument did provide clear insight into the feedback activities of instructors. However, only 1 trained observer did all the scoring. The observational instrument needs to be validated in future DSTC and other complex surgical skills (team) training courses.

For the first time, immediate feedback activities of trainers in an advanced acute trauma simulation course were categorized and investigated. Most feedback activities were instructions on technical skills and had low specificity despite the participation of advanced learners. These study outcomes prompt new research questions: the appropriateness of low-specificity or high-specificity feedback and instruction for optimal learning of advanced learners, the additional effect of non-technical and team skills training, and the effect of ''matching'' instructors' and learners' expectations on teaching and learning outcomes. Ultimately, increased knowledge of feedback in complex surgical skills training will improve course quality, which will be transferred to clinical practice.

REFERENCES

1. Hesketh EA, Laidlaw JM. Developing the teaching instinct: 1. Feedback. Med Teach 2002;24:245–8.
2. van de Ridder JMM, Stokking KM, McGaghie WC, et al. What is feedback in clinical education? Med Educ 2008;42:189–97.
3. Rogers DA, Boehler ML, Schwing CJ, et al. Engaging medical students in the feedback process. Am J Surg 2012;203:21–5.
4. Yarris LM, Linden JA, Hern HG, et al. Attending and resident satisfaction with feedback in the emergency department. Acad Emerg Med 2009;16:S76–81.
5. Richardson BK. Feedback. Acad Emerg Med 2004;11:1238.e1–5.
6. Jaehnig W, Miller ML. Feedback types in programmed instruction: a systematic review. Psychol Rec 2007;57:219–32.
7. Kopp V, Stark R, Fischer MR. Fostering diagnostic knowledge through computer-supported, case- based worked examples: effects of erroneous examples and feedback. Med Educ 2008;42:823–9.
8. Walsh CM, Ling SC, Wang CS, et al. Concurrent versus terminal feedback: it may be better to wait. Acad Med 2009;84:S54–7.
9. Chen X, William RG, Sanfey HA, etal. How do supervising surgeons evaluate guidance provided in the operating room? Am J Surg 2012;203:44–8.
10. Roberts NK, Brenner MJ, Williams RG, et al. Capturing the teachable moment: a grounded theory study of verbal teaching interactions in the operating room. Surgery 2012;151:643–50.
11. Scheeler MC, Bruno K, Grubb E, et al. Generalizing teaching techniques from university to K-12 classrooms: teaching preservice teachers to use what they learn. J Behav Educ 2009;18: 189–210.
12. Goodman JS. The interactive effects of task and external feedback on practice performance and learning. Organ Behav Hum 1998;76: 223–52.
13. Moorthy K, Munz Y, Adams S, et al. A human factors analysis of technical and team skills among surgical trainees during procedural simulation in a simulate operating theatre. Ann Surg 2005;242: 631–9.
14. Yule S, Flin R, Paterson-Brown S, et al. Non-technical skills for surgeons in the operating room: a review of the literature. Surgery 2005;139:140–9.
15. Flin R, Yule S, Paterson-Brown S, et al. Teaching surgeons about nontechnical skills. Surgeon 2007;5:86–9.
16. Crossley J, Marriott J, Purdie H, et al. Prospective observational study to evaluate NOTSS (Non-Technical Skills for Surgeons) for assessing trainees' non-technical performance in the operating theatre. Brit J Surg 2011;98:1010–20.
17. Tapson VF, Karcher RB, Weeks R. Crew resource management and vte prophylaxis in surgery: a quality improvement inititive. Am J Med Qual 2011;26:423–32.
18. Marr M, Hemmert K, Nguyen AH, et al. Team-play in surgical education: a simulation-based study. J Surg Educ 2012;69:63–9.
19. Kohn L, Corrigan J, Donaldson M. To err is human: building a safer health system. Washington, DC: National Academy Press; 1999.
20. Tan ECTH, Egmond, van T, Rots GJM. The emergency war surgery course and the

definitive surgical trauma care course [article in Dutch]. Nederl Mil Geneesk T 2009;62:201–4.
21. Archer JC. State of the science in health professional education: effective feedback. Med Educ 2010;44:101–8.
22. Sweller J, van Merriënboer JJG, Paas F. Cognitive architecture and instructional design. Educ Psychol Rev 1998;10:251–96.
23. van Merriënboer JJG, Sweller J. Cognitive load theory and complex learning: recent developments and future directions. Educ Psychol Rev 2005;17:147–77.
24. Gorma JC, Cooke NJ, Amazeen PG. Training adaptive teams. Hum Factors 2010;52:295–307.
25. Catchpole KR, Dale TJ, Hirst DG, et al. A multicenter trial of aviationstyle training for surgical teams. J Patient Saf 2010;6:180–6.
26. Past G. Styles and strategies of learning. Br J Educ Pscychol 1976;46: 128–48.
27. Hur Y, Kim S. Different outcomes of active and reflective students in problem-based learning. Med Tech 2007;29:e18–21.
28. Vaugh LM, Baker RC. Do different pairings of teaching styles and learning styles make a difference? Preceptor and resident perceptions. Teach Learn Med 2008;20:239–47.
29. Cockerton T, Naz R, Sheppard S. Factorial validity and internal reliability of Honey and Mumfords's learning style questionnaire. Psychol Rep 2002;91:503–19.



CHAPTER 3

Coaching during a trauma surgery team training: perceptions versus structured observations

Alexander Alken[1], Edward Tan[1], Jan-Maarten Luursema[1], Cornelia Fluit[2], Harry van Goor[1]

[1]Department of Surgery
[2]Radboudumc Health Academy

Radboud University Medical Center, Nijmegen, The Netherlands

The American Journal of Surgery 2015; 209 (1): 163-169

ABSTRACT

BACKGROUND: Using the concept perception to quantify coaching skills during surgical training is questionable. This study compared the perceptions reported by the trainees and the faculty members following an emergency surgery team training with structured observations made on the basis of video registrations.

METHODS: For each faculty member, we scored 45 minutes of identical scenarios to enable the quantitative assessment of the use of positive feedback, corrective feedback, as well as instruction and in particular comments containing how and why explanations. We compared the values determined from the video scores with the perceptions reported by the trainees and faculty on questionnaires.

RESULTS: The trainee and faculty ratings for the coaching differed, with trainees generally giving a higher rating. While both the trainees and the faculty gave high ratings for the non-technical skills, corrective and complimenting feedback, and explanations why, the structured video observations showed lower scores in these categories.

CONCLUSION: Both the trainees and the faculty overrated the coaching. Trainee questionnaires and faculty self-reports neither reflected the actual coaching activities nor identified coaching skill deficits.

INTRODUCTION

The operating room (OR) is an important setting for learning surgical skills[1–4] in which the faculty use real-time feedback and instruction to coach trainees.[5–7] The teaching quality of the faculty is of interest to educational researchers and because research into surgical skills encompasses various aspects of teaching and learning, faculty self-reports and trainee questionnaires are considered necessary to examine the faculty's teaching quality and to define areas for improvement.[4]

Faculty self-reports and trainee questionnaires often differ from each other. Compared with trainees, the faculty usually report higher satisfaction with the amount, quality, and usefulness of real-time coaching[2] and feedback.[2,8] Regarding these differences in perceptions, it is not known whether the faculty members overestimate themselves or whether the trainees underestimate the faculty.[4,8] When using research that is based on perceptions to evaluate teaching, it is important to bear in mind that it is not known how close these perceptions reflect reality. Comparisons between perceptions and objective, structured measurements of teaching are required in order to increase the insight into the accuracy of the perceptions of faculty[6] as well as trainees. Studies that have compared perceptions with structured measurements are primarily focused on how self-reporting residents assess their technical skills. Most studies conclude that self-evaluations are unreliable[9,10] and that there are not any or only weak relations between self-evaluations and objective measurements like the Objective Structured Assessment of Technical Skills.[11,12] However, some studies do show stronger relations[13] and an improvement in accuracy with increased experience.[14,15] For non-technical skills, both residents and more experienced surgeons proved not to be accurate when estimating their own performance.

Each year, under the supervision of the Dutch Trauma Society, the Radboud University Medical Center organizes the Definitive Surgical Trauma Care (DSTC) course: an acute surgical trauma team training on live porcine models during which trainees learn damage control techniques while supervised by experienced instructors.[5,16] In the previous 2 editions, the exercises were video and audio taped to permit evaluation of the coaching activities. This provided us with the opportunity to investigate whether trainee and faculty perceptions of coaching agreed with each other. In addition, we could investigate the relationship between these perceptions and the observed findings of coaching on technical and non-technical skills that actually occurred in practice.

METHODS

Participants

The surgical faculty consisted of 11 members. All members were experienced trauma surgeons with 2 to 12 years of teaching experience in trauma surgery. All of them had completed the Advanced Trauma Life Support train-the-trainer course, which includes learning the principles of effective feedback and giving instruction. Fourteen attending surgeons and 7 final-year registrars participated as trainees. All of them worked in hospitals that are involved in resident and/or medical student education. All trainees had limited or no experience in the treatment of the abdominal and thoracic trauma, the topic of this particular DSTC course. During this simulated practical emergency surgery team training, the faculty and the trainees operated in 11 teams. Each operating team consisted of 1 surgical faculty who coached 2 trainees, 1 scrub nurse and 1 member of the anesthesia faculty who coached 1 or 2 anesthesia trainees. Teams were composed in alphabetical order of last names. Only the surgical faculty and surgical trainees participated in this study. Written informed consent documents were obtained before the start of the training. As permitted by the Dutch law, the institutional review board of the Radboud University Medical Center waived the need for formal ethical approval.

Setting

Data were collected during the emergency surgery exercises, which are a part of the 3-day DSTC course.[5,16] They took place at 4 ORs in the university's central animal facility. The learning objective was to apply the principles and techniques of damage control surgery in scenarios using live, anesthetized porcine models with complex and multiple abdominal and thoracic injuries. Trainees were blinded concerning which injuries had been inflicted; both technical and non-technical skills had to be applied within a team setting. Within three and a half hours, the faculty guided the trainees through six acute scenarios. No formal instructions regarding effective coaching were provided. The local animal ethical review board approved this DSTC training.

Outcome measures

The faculty's coaching behavior was recorded and analyzed using an observational coaching instrument (Table 1). Faculty self-reports and trainee questionnaires (Table 2) were taken to compare these 2 perceptions of coaching and to compare perceptions from both the faculty and the trainees with the actual coaching behavior that had been observed during the actual DSTC training practice.

Observational instrument

Coaching activities were recorded real time on audio and video to permit the observation and analysis of the coaching. A 45-minute segment containing 2 scenarios of the approximately 210 minutes recorded per team (6 scenarios) were used for this study: a stab wound to the infra-renal caval vein in the right mid-abdomen and a peri-cardial stab wound to the left lung. These scenarios were selected because they occurred halfway through the training, thereby avoiding missing data by start-up difficulties in the 1st scenario and premature death of the animal in the last scenarios. The 2 scenarios did reflect the acute nature of the injury, the team setting, and the need for technical and non-technical skills, essential components of the complete training.

The faculty's coaching behavior was categorized using a slightly modified version of a previously developed observational instrument[5]: feedback was divided into either corrective or complimenting feedback (Table 1). For each faculty, a trained observer scored the amount of coaching activities and determined the type of reinforcement, ranging from least specific to extremely specific, accompanying the instructions or feedback concerning the team's technical, communication, and team co-operation skills: not reinforced at all (least specific), reinforced with an explanation *how* (moderately specific), an explanation *why* (highly specific), or both *how* and *why* combined (extremely specific). If one sentence addressed more observational categories, then each was scored and independently counted. Potential or ideal opportunities for coaching were not registered. The observer had been trained to observe and analyze recordings during the other 4 DSTC scenarios. In addition, the categorization was ''optimalized'' in a repeated manner during conferences with 2 experienced surgical educators and 2 experienced educationalists until face validity was considered good.

Faculty self-reports and trainee questionnaires

The faculty self-reports and the trainee questionnaires were designed to determine the perceptions of the coaching quality during the complete training. All questionnaire items were repeatedly discussed with the same surgical educators and educationalists until consensus was reached regarding the clarity of the formulated items and the face validity. The faculty self-reports and the trainee questionnaires consisted of 12 items: 9 of which were similar (Table 2, items 1 to 9) and the remaining 3 were comparable (Table 2, items 10 to 12). A 6-point scale was used to rate agreement with each questionnaire item (1 = absolutely not/never, 6 = absolutely yes/all the time) (Table 2).

Each faculty and trainee filled in the questionnaire independently from the other participants; this occurred immediately after the DSTC training had finished. Written instructions on how to fill in the self-reports and questionnaires were provided.

Data analysis

For each faculty member rating the 12 items on the self report, 2 surgical trainees who were coached by this faculty member, rated the same 12 items. To compare faculty with trainee perceptions, the total mean ratings per item of the trainees and that of the faculty members were calculated (Table 2).

Table 1. The modified observational instrument

	COMPLIMENTING FEEDBACK				INSTRUCTIONS			
	Not reinforced at all	*Reinforced with statement how*	*Reinforced with explanation why*	*Reinforced with statement how and explanation why*	*Not reinforced at all*	*Reinforced with statement how*	*Reinforced with explanation why*	*Reinforced with statement how and explanation why*
Technical skills								
Communication skills								
Team cooperation skills								

	CORRECTIVE FEEDBACK			
	Not reinforced at all	*Reinforced with statement how*	*Reinforced with explanation why*	*Reinforced with statement how and explanation why*
Technical skills				
Communication skills				
Team cooperation skills				

It was possible to compare those items (measuring coaching quality) with the observational categories (quantity of coaching behaviors) by combining the observational categories for 8 of the 12 questionnaire items (Table 3). The observed quality of the ''corrective feedback'' and the ''complimenting feedback'' was compared with items 2 and 5. By combining the observational categories ''corrective feedback,'' ''complimenting feedback,'' and ''instructions,'' the observed quantity of coaching on ''technical,'' ''communication,'' and ''team co-operation skills'' was compared with the questionnaire items 7 to 9. The categories ''corrective feedback,'' ''complimenting feedback,'' and ''instructions'' if ''reinforced with only a *how* explanation,'' ''reinforced with only a *why* explanation,'' or ''reinforced with both a *how* and *why* explanation'' were combined and compared with questionnaire item 10. The category ''complimenting feedback'' if ''reinforced with a *why* explanation'' or ''reinforced with both a *how* and *why* explanation'' were combined and the category ''corrective feedback'' when ''reinforced with a *why* explanation'' or ''reinforced with both a *how* and *why* explanation'' were combined and compared with items 11 and 12, respectively.

The mean scores for the faculty were calculated based on the 11 individual faculty scores. Mean differences were calculated between the mean scores of trainees and faculty to determine the between-groups effect. With the calculated mean scores within-group effects were determined within the trainee and the faculty group. Shapiro–Wilk tests were used to determine whether variables were parametric or nonparametric. We used Wilcoxon signed-rank tests for discrete and nonparametric variables and paired-sample *t* tests for continuous and parametric variables. Correlations were calculated by Pearson's *r* (for continuous and parametric variables) and Spearman's ρ (for discrete and nonparametric variables). The comparison between the questionnaire items and the observed data was carried out by determining the correlations and by comparing the within-group effects from the questionnaires with the observed within-group effects. Effect sizes were calculated for *t* tests (Cohen's *d*) and the Wilcoxon signed-rank tests.[17] An effect size of .10 or more is regarded to be small, .30 or more to be medium, and .50 or more to be a large effect. A *P* value of .05 was considered significant. SPSS version 18.0 (SPSS, Inc, Chicago, IL) was used for all analyses.

RESULTS

Eleven self-reports from the faculty members and 20 trainee questionnaires were collected. Two trainees were excluded from the study because they had not returned the informed consent form. The internal consistency of the questionnaires was .86 (Cronbach's α reliability). Table 2 shows the descriptive statistics of the mean ratings for the questionnaire and the observed frequencies for the available matching variables. Overall, the faculty (total questionnaire mean score: 4.5) and trainees (total questionnaire mean score: 5.0) rated satisfaction with the course.

Trainee versus faculty perceptions

Trainee scores were higher for 11 of the 12 items compared with faculty scores (Table 2). The differences were significant for the items corrective feedback (item 2), timing of coaching (item 3), learning as a result of explanations why an action was correct (item 11), and learning as a result of explanations why an action was incorrect (item 12). The faculty rated themselves significantly higher than trainees on giving complimenting feedback (item 5).

The faculty and trainee ratings on coaching for skills related to techniques, communication, and team cooperation were comparable between the 2 groups.

Within-trainee-group effects versus within-faculty-group effects

Trainee ratings for corrective feedback were significantly higher than trainee ratings for complimenting feedback (Table 2, item 2 vs item 5), whereas the faculty ratings did not differ significantly. Trainees' ratings were not significantly different concerning coaching on technical, communication, and team co-operation skills. Similar results were found for the faculty ratings.

Trainee and faculty perceptions versus structured observations

The trainees and the faculty gave high ratings for corrective and complimenting feedback and for explanations as to why an action was correct or incorrect. These ratings, however, did not agree with the observational findings (Table 2): the mean corrective and complimenting feedback was only used 2.1 and 2.3 times, respectively (item 2 and 5); explanations as to why an action was correct had not been used at all (item 11), while explanations why an action was incorrect had only been used a mean of 1.7 times (item 12).

Both the trainees and the faculty showed no significant differences for coaching on technical, communication, and team co-operation skills. This, however, did not agree with the scores of the video and audio observations (Table 2) in which the faculty significantly coached more often on technical skills (mean: 28.2 times; item 7) than on communication (mean: 1.7 times, item 8) or team co-operation skills (mean: .3 times, item 9).

A positive correlation (ρ = .64) was found for the faculty ratings and the observed number of coaching incidents for communication skills.

COMMENTS

The faculty and the trainees both reported an overall high degree of satisfaction for the coaching during practical emergency surgery simulation training. Trainees, however, did rate the timing of that coaching, the corrective feedback, and being able to learn from specified instructions as being higher than the faculty. Both the trainee and the faculty perceptions of the coaching for technical, communication, and team co-operation skills differed from the real-time observation, demonstrating that the coaching on communication and team co-operation skills had been inflated in the perception data.

Table 2. Mean scores per item as perceived by the trainees and faculty (on a 6-point scale) and the available matching frequencies obtained when the coaching was observed during 45 minutes**.**

	Trainees		**Faculty**		**Observed frequencies**	
Questions	Mean	SD	Mean	SD	Mean	SD
Trainees: I perceived the coaching by my faculty trainers as...						
Faculty: I would describe my coaching of the trainees as...						
1. immediately applicable during the operation.	5.2	0.5	4.8	0.9		
2. corrective, focused on what needed to be improved.	5.1*[a]	0.7	4.6	0.7	2.1	2.4
3. well-timed.	5.3*	0.4	4.4	0.5		
4. applicable in next situations.	5.2	0.5	4.7	0.6		
5. complimenting, focused on what went well.	4.4[a]	0.9	5.0#	0.4	2.3	2.2
6. exactly mounted to my personal needs.	4.6	0.7	4.2	0.8		
Trainees: During operating the faculty coached on my...						
Faculty: During operating I coached my trainees on their...						
7. technical skills.	4.8	0.6	4.6	1.2	28.2[b,c]	12.3
8. way of communicating with the team.	4.7	1.1	4.4	1.4	1.7[b,d]	1.7
9. way of cooperating with the team.	4.6	1.1	3.8	1.3	0.3[c,d]	0.6
Trainee: I learned a lot of the...						
Faculty: My coaching contained...						
10. supplementary explanations and tips.	5.4	0.6	5.2	0.4	18.5	12.0
11. reinforcements why an action was correct.	5.2*	0.6	4.7	0.5	0.0	0.0
12. reinforcements why an action was incorrect.	4.9*	0.7	3.9	0.9	1.7	2.5

[a] Trainee within-group effects: corrective feedback higher than complimenting feedback; *mean difference* = 0.7, *P* = .02, *effect size* = 0.5. [b] Effects within structured observations: more coaching on technical skills than communication skills; *mean difference* = 26.5, *t* = 7.66, *P* = .000, Cohen's *d* = 3.0; [c] more coaching on technical skills than team cooperation skills; *mean difference* = 27.9, *P* = .003, Wilcoxon signed rank test, *effect size* = 0.6. [d] more coaching on communication skills than on team cooperation skills; *mean difference* = 1.4, *P* = .02, Wilcoxon signed rank test, *effect size* = 0.5. * Trainees gave higher scores than faculty on: corrective feedback (*mean difference* = 0.6, *P* = .04, *effect size* = 0.4); timing of coaching (*mean difference* = 0.6, *P* = .01, *effect size* = 0.6); teaching following explanations why an action was correct (*mean difference* = 0.5, *P* = .046, *effect size* = 0.4); teaching following explanations why an action was incorrect (*mean difference* = 1.0, *P* = .03, *effect size* = 0.5). # The Faculty gave higher ratings than the trainees on: complimenting feedback (*mean difference* = 0.6, *P* = .03, *effect size* = 0.5).

Table 3. The correspondences between the coaching questionnaires and the categories of the observational instrument.

Questionnaire item		Represented by the observed amount within the observational categories:
2. Corrective coaching	→	Corrective feedback; - *either: not reinforced at all, reinforced with statement how, reinforced with explanation why, reinforced with statement how and explanation why;* - either: pertaining to technical, communication and team co-operation skills.
5. Complimenting coaching	→	Complimenting feedback; - *either: not reinforced at all, reinforced with statement how, reinforced with explanation why, reinforced with statement how and explanation why;* - either: pertaining to technical, communication or team co-operation skills.
7. Coaching on technical skills	→	Complimenting feedback + corrective feedback + instructions; - *either: not reinforced at all, reinforced with statement how, reinforced with explanation why, reinforced with statement how and explanation why;* - only: pertaining to technical skills.
8. Coaching on communication skills	→	Complimenting feedback + corrective feedback + instructions; - *either: not reinforced at all, reinforced with statement how, reinforced with explanation why, reinforced with statement how and explanation why;* - only: pertaining to communication skills.
9. Coaching on team co-operation skills	→	Complimenting feedback + corrective feedback + instructions; - *either: not reinforced at all, reinforced with statement how, reinforced with explanation why, reinforced with statement how and explanation why;* - only: pertaining to team co-operation skills.
10. Coaching with supplementary explanations and tips	→	Complimenting feedback + corrective feedback + instructions; - *only: reinforced with statement how, reinforced with explanation why, reinforced with statement how and explanation why;* - either: pertaining to technical, communication or team co-operation skills.
11. Coaching with reinforcements why correct	→	Complimenting feedback; - *only: reinforced with explanation why, reinforced with statement how and explanation why;* - either: pertaining to technical, communication or team co-operation skills.
12. Coaching with reinforcements why incorrect	→	Correcting feedback; - *only: reinforced with explanation why, reinforced with statement how and explanation why;* - either: pertaining to technical, communication or team co-operation skills.

In contrast to most studies,[2,8] trainees gave the faculty a higher rating than the faculty gave themselves. All trainees were attending surgeons or final-year registrars who worked in hospitals in which they taught junior residents or medical students. Possibly their own teaching experience made them understand the complexity of OR teaching and made them more lenient in rating the faculty's teaching.[18] Furthermore, senior trainees prefer training on live animal models above other simulations because of the realistic circumstances. This preference seems to be independent of the didactic course design or educational value.[19]

Our real-time observational data indicated that the faculty overrated their coaching abilities on non-technical skills and the type and explanatory character of the coaching. The faculty may have rated their potential level instead of rating to the level at which they actually performed.[15,20] It is also possible that the faculty lack appropriate insight[15] into their coaching activities and thus overestimate their capabilities. However, Evans et al[20] stated that the pressure to perform well and the need to impress is a more plausible explanation.

This study's findings raise the question whether, in general, individuals are able to evaluate themselves with sufficient accuracy. The process of self-evaluation is complex and highly susceptible to contextual factors, therefore resulting in variable and inconsistent self-evaluations.[10] Nevertheless, it is important to strive for accurate self-reports to establish the potential to learn and develop professional expertise.[13,15] Video-based self-reflection has been shown to be an effective method to improve one's ability to determine his own strengths and weaknesses.[15] It may also be an effective method to improve one's future teaching. In other research, however, external evaluations have been recommended to improve the reliability in identifying one's strengths, weaknesses, and areas requiring improvement.[10] The finding that those who have the lowest performance most often overrate themselves[9,18] supports the need to use external evaluations. That no differences in overrating were apparent in this study is most likely because of the uniformly high rates with small standard deviations.

Evaluation by the trainees is considered necessary to examine the faculty's teaching abilities and to identify those areas requiring improvement.[4] We found that our trainees considered the faculty to apply important coaching activities that had rarely been observed during the video and audio analysis. For future evaluation, it seems necessary to first train trainees in the criteria required to allow them to ''objectively'' evaluate courses and the faculty's teaching abilities, especially when trainees and faculty may be involved in long-term training programs. Equipping trainees to be able to recognize effective and ineffective teaching most probably will improve future evaluations[4,8] and will make it possible to identify areas that require improvement.

Although the need was stressed in 2002,[18] more than a decade ago, valid tools to observe and analyze the quality of coaching during complex practical training are still not available. We have begun the process of validating our observational instrument in the DSTC and other emergency surgery courses. Our questionnaires were developed in close consultation with experts in medical and surgical education and have showed good internal consistency. The results obtained indicated that the perceptions were not in line with what we actually observed. While we selected 45 minutes from the operating scenarios, the trainees and faculty based their perceptions on the whole training. Therefore it remains possible that during other parts of the training the faculty did actually provide more coaching on non-technical skills. In our research, we compared quantitative observations with qualitative questionnaire outcomes of coaching perceptions. Even though a comparison has been made between quantitative and qualitative data, we think it is justified to conclude that these comparisons pointed out that the coaching and training would have been more effective if non-technical skills and detailed explanations would have been provided more often.

We compared the faculty's self-perceptions with the trainee's perceptions obtained by questionnaires to structured video observations of a practical surgical simulation team training for emergency trauma procedures to analyze the teaching as it occurred. Both the trainees and the faculty overrated the coaching activities, with trainees' ratings being higher than faculty's. This indicates that, at least in such emergency surgery courses, the use of trainee questionnaires and faculty self-reports will not provide adequate data to permit an analysis of the quantity and quality of the faculty's coaching and to identify coaching skills deficits.

REFERENCES

1. Reznick RK, MacRae H. Teaching surgical skills–changes in the wind. N Engl J Med 2006;355:2664–9.
2. Levinson KL, Barlin JN, Altman K, et al. Disparity between resident and attending physician perceptions of intraoperative supervision and education. J Grad Med Educ 2010;2:31–6.
3. Iwaszkiewicz M, Darosa DA, Risucci DA. Efforts to enhance operating room teaching. J Surg Educ 2008;65:436–40.
4. Butvidas LD, Anderson CI, Balogh D, et al. Disparities between resident and attending surgeon perceptions of intraoperative teaching. Am J Surg 2011;201:385–9; discussion 389.
5. Alken A, Tan E, Luursema JM, et al. Feedback activities of instructors during a trauma surgery course. Am J Surg 2013;206: 599–604.
6. Chen XP, Williams RG, Sanfey HA, et al. How do supervising surgeons evaluate guidance provided in the operating room? Am J Surg 2012;203:44–8.
7. Hauge LS, Wanzek JA, Godellas C. The reliability of an instrument for identifying and quantifying surgeons' teaching in the operating room. Am J Surg 2001;181:333–7.
8. Yarris LM, Linden JA, Gene Hern H, et al. Attending and resident satisfaction with feedback in the emergency department. Acad Emerg Med 2009;16(Suppl 2):S76–81.
9. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician selfassessment compared with observed measures of competence: a systematic review. JAMA 2006;296:1094–102.
10. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Acad Med 2005;80(10 Suppl): S46–54.
11. Pandey VA, Wolfe JH, Black SA, et al. Self-assessment of technical skill in surgery: the need for expert feedback. Ann R Coll Surg Engl 2008;90:286–90.
12. Brewster LP, Risucci DA, Joehl RJ, et al. Comparison of resident selfassessments with trained faculty and standardized patient assessments of clinical and technical skills in a structured educational module. Am J Surg 2008;195:1–4.
13. Arora S, Miskovic D, Hull L, et al. Self vs expert assessment of technical and non-technical skills in high fidelity simulation. Am J Surg 2011;202:500–6.
14. Ward M, MacRae H, Schlachta C, et al. Resident self-assessment of operative performance. Am J Surg 2003;185:521–4.
15. Moorthy K, Munz Y, Adams S, et al. Self-assessment of performance among surgical trainees during simulated procedures in a simulated operating theater. Am J Surg 2006;192:114–8.
16. Tan ECTH, Egmond van T, Rots GJM. The emergency war surgery course and the definitive surgical trauma care course. Nederl Mil Geneesk T 2009;62:201–4. summary in English.
17. Field A. Non-parametric tests. In: Field A, ed, editor. Discovering Statistics Using SPSS. 3rd ed. London: Sage; 2009.
18. Claridge JA, Calland JF, Chandrasekhara V, et al. Comparing resident measurements to

attending surgeon self-perceptions of surgical educators. Am J Surg 2003;185:323–7.
19. Palter VN, Orzech N, Aggarwal R, et al. Resident perceptions of advanced laparoscopic skills training. Surg Endosc 2010;24: 2830–4.
20. Evans AW, Leeson RM, Newton-John TR. The influence of self deception and impression management on surgeons' self-assessment scores. Br Dent J 2005;198:765–9.

3

CHAPTER 4

Integrating technical and non-technical skills coaching in an acute trauma surgery team training: Is it too much?

Alexander Alken[1], Jan-Maarten Luursema[1], Mariska Weenk[1], Simon Yauw[1], Cornelia Fluit[2], Harry van Goor[1]

[1]Department of Surgery
[2]Radboudumc Health Academy

Radboud University Medical Center, Nijmegen, The Netherlands

The American Journal of Surgery 2018; 216 (2): 369-374

ABSTRACT

OBJECTIVE: Research on effective integration of technical and non-technical skills in surgery team training is sparse. In a previous study we found that surgical teachers predominantly coached on technical and hardly on non-technical skills during the Definitive Surgical and Anesthetic Trauma Care (DSATC) integrated acute trauma surgery team training. This study aims to investigate whether the priming of teachers could increase the amount of non-technical skills coaching during such a training.

DESIGN: Coaching activities of 12 surgical teachers were recorded on audio and video. Six teachers were primed on non-technical skills coaching prior to the training. Six others received no priming and served as controls. Blind observers reviewed the recordings of 2 training scenarios and scored whether the observed behaviors were directed on technical or non-technical skills. We compared the frequency of the non-technical skills coaching between the primed and the non-primed teachers and analyzed for differences according to the trainees' level of experience.

SETTING: Surgical teachers coached trainees during the highly realistic DSATC integrated acute trauma surgery team training. Trainees performed damage control surgery in operating teams on anesthetized porcine models during 6 training scenario's.

PARTICIPANTS: Twelve experienced surgical teachers participated in this study.

RESULTS: Coaching on non-technical skills was limited to about 5%. The primed teachers did not coach more often on non-technical skills than the non-primed teachers. We found no differences in the frequency of non-technical skills coaching based on the trainees' level of experience.

CONCLUSION: Priming experienced surgical teachers does not increase the coaching on non-technical skills. The current DSATC acute trauma surgery team training seems too complex for integrating training on technical and non-technical skills.

INTRODUCTION

Shortcomings in non-technical skills are important contributors to errors and adverse events in the operating room.[1-6] Training improves non-technical skills in the operating room and reduces errors,[6-8] patient morbidity[9] and mortality.[10] As a result, non-technical skills training is becoming an essential part of surgical training.[7]

Training can be specifically designed for non-technical skills development.[6] Although effective, other researchers promote integrated skills training in which technical skills and non-technical skills are trained concurrently[11-15] in order to enhance the transfer of non-technical skills to the operating room.[14,16] However, our group found that integrating technical and non-technical skills in surgical training resulted in a strong focus on technical skills.[17,18] There is sparse evidence for the integration of technical and non-technical skills training in surgical specialties[11,12,15] which makes it difficult to design effective and evidence based integrated skills training courses.

At our academic institute we annually offer an integrated acute trauma surgery team training, which is part of the national Definitive Surgical and Anesthetic Trauma Care (DSATC) course. In this highly complex training setting we found that the coaching on non-technical skills was not in proportion to the coaching on technical skills.[17,18] To gain more insight in the possibilities to increase the amount and quality of non-technical skills teaching next to technical skills teaching, we investigated whether the priming of surgical teachers could increase the amount of non-technical skills coaching. We expect a priming effect in which behavior (in this case non-technical skills coaching) can be evoked, even unconsciously, due to previously offered cues.[19] Theoretical knowledge of non-technical skills and a group discussion on how to coach non-technical skills were the core activities of this priming as research showed that surgical teachers perceive this as useful.[20] In this comparative study we hypothesized that primed teachers would coach more frequently on non-technical skills than non-primed teachers.

METHODS

Participants

We analyzed the coaching of 12 experienced trauma surgeons with 2-32 years of teaching experience who taught in previous DSATC courses. All completed the Advanced Trauma Life Support (ATLS) student course on technical and non-technical aspects of treating critically injured patients, and the ATLS instructor training on teaching in emergency simulation. Twenty-four surgical trainees participated: 8 final year residents (6 year program), 8 junior surgeons (0-5 years of experience) and 8 senior surgeons (6-25 years of experience). Participation is voluntary for general and resident surgeons. Trauma surgeons are required to participate every 4 years.[21] Twelve teams were composed, each containing 1 surgical teacher, 2 surgical trainees, 1 experienced scrub nurse, 1 experienced anesthesiologist teacher and 1 or 2 anesthesia trainees. Participants worked in different hospitals all over the country and were unlikely to have previously collaborated. Surgical trainees were allocated into teams according to the alphabetical order of their surname. Only the surgical teachers were subject to this study. Voluntary written informed consents were collected. The local review board waived the need for formal ethical approval.

Course description

The DSATC course is a three-day damage control surgery course, combining the original Definitive Surgical Trauma Care (DSTC)[22] and the Definitive Anesthetic Trauma Care (DATC) course.[23] Each day starts with 3.5 h of lectures followed by a 3.5-h workshop. All teachers and trainees attend the lectures on day 1, 2 and 3. All lectures are about technical skills, except for one 30-min lecture on non-technical skills on day 1. Every day, 4 teachers (12 individuals in 3 days) and 8 trainees (24 in 3 days) participate in the 3.5-h integrated acute trauma surgery team training on anesthetized porcine models in 4 operating rooms at the animal facility. This workshop is the focus of our study. Each teacher coaches 2 surgical trainees through 6 scenarios, each containing one or two complex, life-threatening abdominal or thoracic injuries which are not common practice. Trainees are blinded for the injuries inflicted by the teachers. The primary aims are to teach surgical technical skills (damage control techniques) and multidisciplinary non-technical communication and team skills in a team setting.[21,23] The DSATC training is approved by the animal ethical review board.

Intervention and control group

The surgical teachers were assigned to an intervention (N = 6) and control group (N = 6). The intervention group coached trainees on day 1 (N = 4) and 2 (N = 2), the control group on day 2 (N = 2) and 3 (N = 4). Both groups attended the general program containing the 30-min lecture and a spoken reminder (Table 1). In addition, the intervention group participated in a 20-min priming provided by one educationalist/doctor of medicine and one educationalist, both specialized in medical and surgical teaching (AA and CF) (Table 2). The purpose was to define non-ambiguous definitions and examples of communication and team skills (e.g. *timely inform the anesthetist about bleedings*) since a similar intervention was considered helpful in non-technical skills teaching.[20] Opportunities for teaching during the training were explicitly discussed. A 20-min priming was deemed sufficient because we expected a priming effect and all surgical teachers were experienced, taught in previous DSATC integrated team training courses and received training in non-technical skills. The intervention group was told not to share any information with the control group.

Data collection

Training was recorded on audio and video. Scenarios 1 and 3 were transcribed, reviewed and coaching was scored by 2 observers with an earlier developed instrument.[18] We selected scenario 1 (a liver bleeding and bleeding of the groin) because it was the first scenario after the priming and reminder, and scenario 3 (a second liver bleeding) to gain an impression of how the non-technical skills coaching developed over time. Non-technical skills were equally important in all scenarios.

For each non-technical utterance we analyzed to which trainee it was addressed and what his/her level of experience was. We collected the same amount of technical utterances. If 2 trainees were addressed concurrently (e.g. general explanations) or if it was unclear which trainee was addressed (e.g. blocked faces on camera) one utterance was scored twice as being directed to trainee A and B.

4

Table 1. Activities of the intervention and control group.

Timeline	Event	Intervention group	Control group
Evening before the start of the DSATC course	Additional meeting	20-min priming on non-technical skills coaching by 2 educationalists (Table 2)	No intervention
Day 1 of the DSATC course	General program: 3.5 h of lectures	30-min lecture on aviation style communication and team skills; provided by an air craft pilot instructor	30-min lecture on aviation style communication and team skills; provided by an air craft pilot instructor
Day 1-3 of the DSATC course	General program: briefing right before starting the integrated acute trauma surgery team training	Spoken reminder to pay sufficient attention to teaching team and communication skills; provided by the local course organizer	Spoken reminder to pay sufficient attention to teaching team and communication skills; provided by the local course organizer

Observer training

Two blinded surgical residents were trained by researcher AA in scoring coaching with an observational instrument[18] using recordings of a previous DSATC course, until an inter-rater reliability of Cohen's $K = 0.7$ was reached. The observational categories for teaching were: (1) "*Coaching on non-technical skills*" (the sum of all instances of '*Positive feedback*', '*Corrective feedback*' and '*Instruction*' directed on '*Communication skills*' and '*Team collaboration skills*'); (2) "*Coaching on technical skills*" (the sum of all '*Positive feedback*', '*Corrective feedback*' and '*Instruction*' directed on '*Technical skills*');[18] and (3) '*Other coaching*', referring to relevant coaching which could not be categorized in '*Non-technical skills*' or '*Technical skills*' (e.g. introducing the patient case, asking questions to check on the progress).

Table 2. Content of the priming provided to the intervention group.

Part		Focus	Time
1	Last year's research to coaching during the DSATC	a. Presenting our findings which showed a predominant attention on technical skills teaching (93%) in comparison with the teaching of communication (6%) and team cooperation skills (1%). b. Discussing our findings with the teachers of the intervention group.	5 min
2	Defining non-technical skills	a. Presenting the evidence-based importance of non-technical skills in the operating room. b. Presenting definitions of communication and team co-operation skills and provide exemplar behaviors.	5 min
3	Group discussion	a. Presenting a tips-and-tricks summary on how to recognize and create moments for non-technical skills teaching and increase the attention for non-technical skills during training. b. Group discussion on how teachers think they can increase the coaching on non-technical skills and improve the feedback and instruction on non-technical skills during the training.	10 min

Data analysis

"*Coaching on non-technical skills*" and "*Coaching on technical skills*" was merged with all "*Other coaching*" to quantify the amount of "*Total coaching*". To correct for talkativeness, the relative amount of "*Coaching on non-technical skills*" was calculated as a percentage of the "*Total coaching*" per teacher per scenario (Table 3). The means for the intervention and control group were calculated and compared for scenario 1, scenario 3, 1 and 3 combined and the difference in frequency between 3 and 1. Mann Whitney U-tests were used to analyze differences.

Per teacher per scenario we collected all non-technical skills coaching utterances and the same amount of technical skills coaching utterances. Since the non-technical skills coaching utterances were distributed throughout the training, we divided the total amount of technical skills utterances (e.g. 91) by the total amount of non-technical skills utterances (e.g. 9) and collected every *x*'th (e.g. 91/9 = 10.11; every 10th) technical skills coaching utterance made by that teacher to reflect similar distributions. We then analyzed for all coaching utterances (100%) per trainee level (resident, junior surgeon or senior surgeon trainee) how much was technical and non-technical. We used the Kruskal-Wallis test to analyze differences between the percentage of non-technical skills coaching within the three trainee groups. SPSS version 21.0 (SPSS, Inc, Chicago, IL) was used for all analyses. A *P*-value less than 0.05 was considered significant.

Table 3. Observed frequencies of coaching and the relative amount of non-technical skills coaching in a percentage of the total amount of coaching for the intervention and the control group.

		Intervention group						Control group			
		Observed frequency		Relative amount of NTS coaching[a]				Observed frequency		Relative amount of NTS coaching[a]	
Instructor 1	**Coaching on NTS**	**9**	**0**	**7.96%**	**0.00%**	**Instructor 7**	**Coaching on NTS**	**9**	**1**	**6.62%**	**1.14%**
Intervention group	Coaching on TS	78	30			*Control group*	Coaching on TS	96	79		
	Other coaching	26	12				Other coaching	31	8		
	Total coaching	113	42				Total coaching	136	88		
Instructor 2	**Coaching on NTS**	**5**	**5**	**3.82%**	**7.25%**	**Instructor 8**	**Coaching on NTS**	**0**	**1**	**0.00%**	**3.03%**
Intervention group	Coaching on TS	100	54			*Control group*	Coaching on TS	47	30		
	Other coaching	26	10				Other coaching	10	2		
	Total coaching	131	69				Total coaching	57	33		
Instructor 3	**Coaching on NTS**	**2**	**2**	**2.60%**	**5.00%**	**Instructor 9**	**Coaching on NTS**	**11**	**0**	**14.86%**	**0.00%**
Intervention group	Coaching on TS	50	27			*Control group*	Coaching on TS	36	18		
	Other coaching	25	11				Other coaching	27	12		
	Total coaching	77	40				Total coaching	74	30		
Instructor 4	**Coaching on NTS**	**6**	**2**	**6.90%**	**4.76%**	**Instructor 10**	**Coaching on NTS**	**9**	**1**	**7.83%**	**3.45%**
Intervention group	Coaching on TS	64	35			*Control group*	Coaching on TS	84	18		
	Other coaching	17	5				Other coaching	22	10		
	Total *coaching*	87	42				Total *coaching*	115	29		
Instructor 5	**Coaching on NTS**	**9**	**1**	**10.23%**	**2.56%**	**Instructor 11**	**Coaching on NTS**	**4**	**1**	**7.69%**	**2.17%**
Intervention group	Coaching on TS	64	32			*Control group*	Coaching on TS	27	29		
	Other coaching	15	6				Other coaching	21	16		
	Total coaching	88	39				Total coaching	52	46		
Instructor 6	**Coaching on NTS**	**2**	**0**	**2.17%**	**0.00%**	**Instructor 12**	**Coaching on NTS**	**6**	**1**	**4.20%**	**2.33%**
Intervention group	Coaching on TS	83	48			*Control group*	Coaching on TS	112	35		
	Other coaching	7	0				Other coaching	25	7		
	Total coaching	92	48				Total coaching	143	43		
Mean relative amount of NTS coaching per scenario				**5.61%**	**3.26%**	Mean relative amount of NTS coaching per scenario				**6.87%**	**2.02%**
				SD = 3.24	*SD = 2.93*					*SD = 4.90*	*SD = 1.27*
Mean relative amount of NTS coaching for scenario 1 and 3 combined				**4.44%** *SD = 3.19*		Mean relative amount of NTS coaching for scenario 1 and 3 combined				**4.44%** *SD = 4.25*	

TS = Technical skills; NTS = Non-technical skills.

[a] Formula used to calculate the relative amount of NTS coaching: *Coaching on NTS/Total coaching * 100%*

RESULTS

All 12 surgical teachers were male. Twenty-three surgical trainees were male, one trainee was female. Data of 6 teachers in the intervention group (with a mean teaching experience of 17 years) and 6 teachers in the control group (with a mean teaching experience of 16 years) was analyzed. A total of 1714 utterances were recorded: 868 in the intervention group and 846 in the control group. Overall, the intervention group coached 665 times on technical skills, 43 times on non-technical skills and applied other coaching 160 times. The control group coached 611 times on technical skills, 44 times on non-technical skills and applied other coaching 191 times.

Table 4. The mean relative amount of non-technical skills coaching in percentage of the total amount of coaching applied by the surgical teachers of each group.

		Coaching on NTS/Total coaching * 100%	
		Intervention group	**Control group**
Scenario 1	Mean	5.61%	6.87%[a]
	SD	*3.24*	*4.90*
Scenario 3	Mean	3.26%	2.02%[b]
	SD	*2.93*	*1.27*
Scenario 1 and 3	Mean	4.44%	4.44%[c]
	SD	*3.19*	*4.25*

NTS = Non-technical skills

[a] Mean difference intervention and control group = 1.26; P = .818, Mann Whitney U Test.

[b] Mean difference intervention and control group = 1.24; P = .485, Mann Whitney U Test.

[c] Mean difference intervention and control group = 0.00; P = .713, Mann Whitney U Test.

The relative amount of non-technical skills coaching by the intervention group accounted for 0.00%-10.23% of the total coaching, depending on teacher and scenario (Table 3). The intervention group directed a mean relative amount of 5.61% of their coaching on non-technical skills during scenario 1, a mean of 3.26% during scenario 3 and a mean of 4.44% during scenario 1 and 3 combined. The relative amount of non-technical skills coaching by the control group accounted for 0.00%-14.86% of the total coaching, depending on teacher and scenario (Table 3). The control group directed a mean relative amount of 6.87% of their coaching on non-technical skills during scenario 1, a mean of 2.02% during scenario 3 and a mean of 4.44% during scenario 1 and 3 combined. We found no statistically significant differences on non-technical skills coaching between the intervention and the control group (Table 4). When comparing the non-technical skills coaching in scenario 3 to scenario 1 we found a significant

decrease in the amount of non-technical skills coaching for the control group compared to the intervention group (Table 5).
There were 4 resident-junior surgeon trainee pairs, 4 resident-senior surgeon trainee pairs, 2 junior-senior surgeon trainee pairs, 1 junior-junior surgeon trainee pair and 1 senior-senior surgeon trainee pair. We analyzed for 87 non-technical skills coaching utterances (the total amount of non-technical skills coaching applied by the intervention and control group together) and for 87 technical skills coaching utterances to which trainee level they were directed: 82 utterances were directed to residents (40 (48.78%) on technical skills and 42 (51.22%) on non-technical skills); 83 to junior surgeons (41 (49.40%) on technical skills and 42 (50.60%) on non-technical skills) and 84 to senior surgeons (42 (50%) on technical skills and 42 (50%) on non-technical skills (174 utterances in total; 75 utterances counted twice because they were directed to two trainees of different levels simultaneously). There were no statistically significant differences in percentages of non-technical skills coaching with regard to the trainee levels (P = 0.368, Kruskal-Wallis test).

4

Table 5. The differences in the mean relative amount of non-technical skills coaching in percentage of the total amount of coaching applied by the surgical teachers of each group.

	Scenario 3 – Scenario 1
Intervention group	3.26% - 5.61% = -2.35%[a]
Control group	2.02% - 6.87% = -4.85%[b]

NTS = Non-technical skills.
[a] 3.26% - 5.61% = -2.35% (Intervention group); P = .310, Mann Whitney U Test.
[b] 2.02% - 6.87% = -4.85% (Control group); P = .041, Mann Whitney U Test.

DISCUSSION

A 20-min priming of experienced surgical teachers on non-technical skills coaching with a preoperative reminder does not increase non-technical skills coaching during the DSATC integrated acute trauma surgery team training. Primed teachers might coach more consistently on non-technical skills throughout the training than non-primed teachers, however, this effect is only minor. The trainees' level of experience does not trigger different coaching behavior.
The priming might not have been sufficient to strengthen coaching behavior of the intervention group in addition to the 30min non-technical skills lecture and the reminder to teach non-technical skills, which was provided to all teachers. We doubt, however, whether a longer and more intense priming would have increased the coaching on non-technical skills in this highly complex simulation training. All teachers were experienced trauma surgeons who successfully completed instructor training programs, had extensive teaching experience and were trained

in aspects of non-technical skills. In that perspective the 20-min priming on non-technical skills could be considered a refresher.
However, during the DSATC integrated acute trauma surgery team training trainees learn and train unfamiliar and technically challenging techniques on anesthetized porcine models with severe bleedings under real time pressure. Such circumstances are known for the risk to cause acute stress,[24-26] cognitive overload and the inability to split attention.[27] Simulation training should occur in a safe training environment in which trainees can deliberately practice their skills and teachers can focus on teaching as their primary objective instead of patient care.[28] We feel that this criterion is not met in the DSATC integrated acute trauma surgery team training. More specifically, based on the findings of our current study and consistent findings of previous studies on non-technical skills coaching[17,18] we believe that the training environment is too realistic, too hectic and too unstable to enable teachers to coach on non-technical and technical skills concurrently. The complexity and the risk of losing the "patient" may prevent teachers from coaching on non-technical skills, even if they are appropriately certified, trained and primed prior to the course. For trainees the training may be too complex to pay attention to both technical and non-technical skills. Our previous studies also showed that coaching (both technical and non-technical) was rarely specific and explained in detail.[17,18]
We analyzed for differences in coaching behavior related to the trainees' level of experience, assuming that stress and cognitive overload is less in experienced trainees which would result in more opportunities for teachers to address non-technical skills. The finding that the coaching behavior of teachers was comparable in less and more experienced trainees challenges this theory and supports our feeling that the training environment is too complex for training technical and non-technical skills concurrently. However, the levels of most trainees within each pair differed which might have obscured differences in coaching behavior. It is questionable whether it is possible at all to coach a single trainee if trainees operate together intensely.
According to the manual, the course objectives of the original DSTC encompass the development of surgical decision-making and techniques to manage the critically injured patient.[22] The use of anesthetized porcine models is advocated to teach and train surgical psychomotor skills and techniques to control bleedings and to save organs. It should be "a challenge to both the veterinary anesthetist and the surgeon to maintain a viable animal".[22] By combining the DSTC with the DATC an anesthetic training team is added to enable and emphasize training in effective multidisciplinary team work strategies mimicking the clinical situation.[21,23] Regardless whether a veterinary anesthetist (DSTC) or a patient anesthetic team (DATC) is involved, communication and collaboration skills are essential for performing a successful operation. Remarkably, in comparison with the technical objectives, the non-technical objectives are described in far less detail in the manuals. If technical and non-technical skills are continued to be trained concurrently in such a complex environment we suggest a firm embedding of non-technical skills coaching in the design of each scenario. For instance, teams must achieve haemostasis

in scenario 1 (when gaining control of a liver bleeding and bleeding of the groin) and scenario 3 (when gaining control of a second liver bleeding) and have a one or 2-min waiting time before the haemostatic agent becomes effective. These moments can be purposely dedicated to pause and reflect on non-technical skills coaching before continuing the operation. Reflective debriefings show benefits in teaching and training non-technical skills,[29] also within a training.[30] An alternative DSATC integrated acute trauma surgery team training may be possible by applying principles of the four component instructional design (4 C/ID) model[31] in a training environment known as the virtual operating room.[32] The 4 C/ID model is an effective strategy for teaching complex skills.[16,31,33] One of its key elements is to teach trainees whole-task procedures in authentic learning environments, covering all necessary competencies from the onset of training, and gradually increase the task complexity. The virtual operating room is an authentic learning environment in which technical and non-technical skills can be trained in a team setting, with trainees performing surgical procedures on virtual reality simulators, bench top models and/or human patient simulators managed from a separate control room.[32] The condition of the "patient" is more controllable than an anesthetized animal model which enables teachers to pause the condition of the patient, even in acute situations, and to elaborately coach trainees.[34,35] When applying this model to the DSATC integrated acute trauma surgery team training, we propose to implement multiple, distributed training sessions, starting with teaching the necessary technical and non-technical skills in simple trauma procedures in uncomplicated simulated patients, and gradually progressing to more complex procedures in more complicated simulated patients. The training on an anesthetized live animal could be considered as an endpoint of this training series. We hypothesize this would be an effective method to effectively integrate and train technical and non-technical skills and to establish transfer to real practice. Future research on its effectiveness is needed.

Our study has a few limitations. First, the sample size of teachers is small with the risk of selection bias. Second, the about 5% of coaching that was directed on non-technical skills was low, which reduced the chance to obtain significant differences. Third, the focus of our research was on changing teachers' non-technical skills coaching behavior. We did not investigate and did not control for variables in other team members which could have influenced our findings. Team members could have had some prior experience in working together. However, this is highly unlikely since participants came from all over the country. We did not investigate prior experiences in and perceptions towards non-technical skills. Practical surgical team experience was present in all team members which would have made a lack of understanding unlikely. However, there could still have been a lack of awareness of the importance of non-technical skills among team members which might have hampered the coaching on non-technical skills. Future research should address these topics. Lastly, we did not looked at total and missed opportunities for technical and non-technical coaching which would have quantified the potential for coaching. Proper scoring of opportunities to coach is difficult and subjective with expected high inter-rater variation.

The findings of this study provide a valuable contribution to the debate on how to integrate technical and non-technical skills teaching in simulation team training and the future design of complex integrated skills training courses. As a possible solution for successful concurrent technical and non-technical skills training we advocate a less hectic and more controllable training environment with multiple pause and reflect procedures dedicated to non-technical skills. We advise to reconsider non-technical skills' teaching objectives in highly complex and realistic simulations such as the DSATC acute trauma surgery team training. Future research should be directed on how to successfully coach and train technical and non-technical skills concurrently, investigating the learning progress of trainees, analyze the possible effects of these pause and reflect procedures and the transfer of non-technical skills to the operating room after such integrated skills training programs.

REFERENCES

1. Greenberg CC, Regenbogen SE, Studdert DM, et al. Patterns of communication breakdowns resulting in injury to surgical patients. J Am Coll Surg. 2007;204: 533-540.
2. Kohn LT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System. Washington (DC): National Academies Press (US); 2000. Copyright 2000 by the National Academy of Sciences. All rights reserved.
3. Lingard L, Espin S, Whyte S, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. Qual Saf health care. 2004;13:330-334.
4. Mazzocco K, Petitti DB, Fong KT, et al. Surgical team behaviors and patient outcomes. Am J Surg. 2009;197:678-685.
5. Siu J, Maran N, Paterson-Brown S. Observation of behavioural markers of nontechnical skills in the operating room and their relationship to intra-operative incidents. Surg. 2016;14:119-128.
6. Weaver S, Dy S, Rosen M. Team-training in healthcare: a narrative synthesis of the literature. BMJ Qual Saf. 2014;23:359-372.
7. Dedy NJ, Bonrath EM, Zevin B, Grantcharov TP. Teaching nontechnical skills in surgical residency: a systematic review of current approaches and outcomes. Surgery. 2013;154:1000-1008.
8. McCulloch P, Rathbone J, Catchpole K. Interventions to improve teamwork and communications among healthcare staff. Br J Surg. 2011;98:469-479.
9. Young-Xu Y, Neily J, Mills PD, et al. Association between implementation of a medical team training program and surgical morbidity. Arch Surg (Chicago, Ill 1960). 2011;146:1368-1373.
10. Neily J, Mills PD, Young-Xu Y, et al. ASsociation between implementation of a medical team training program and surgical mortality. JAMA. 2010;304: 1693-1700.
11. Brunckhorst O, Shahid S, Aydin A, et al. Simulation-based ureteroscopy skills training curriculum with integration of technical and non-technical skills: a randomised controlled trial. Surg Endosc. 2015;29:2728-2735.
12. Dawe SR, Pena GN, Windsor JA, et al. Systematic review of skills transfer after surgical simulation-based training. Br J Surg. 2014;101:1063-1076.
13. Whittaker G, Abboudi H, Khan MS, Dasgupta P, Ahmed K. Teamwork assessment tools in modern surgical practice: a systematic review. Surg Res Pract. 2015;2015, 494827.
14. Azuara-Blanco A, Reddy A, Wilkinson G, Flin R. Safe eye surgery: non-technical aspects. Eye Lond Engl. 2011;25:1109-1111.
15. Brunckhorst O, Shahid S, Aydin A, et al. The relationship between technical and nontechnical skills within a simulation-based ureteroscopy training environment. J Surg Educ. 2015;72:1039-1044.
16. Vandewaetere M, Manhaeve D, Aertgeerts B, Clarebout G, Van Merrienboer JJ, Roex A. 4C/ID in medical education: how to design an educational program based on whole-task learning: AMEE Guide No. 93. Med Teach. 2015;37:4-20.
17. Alken A, Tan E, Luursema JM, Fluit C, van Goor H. Feedback activities of instructors during a trauma surgery course. Am J Surg. 2013;206:599-604.

18. Alken A, Tan E, Luursema JM, Fluit C, van Goor H. Coaching during a trauma surgery team training: perceptions versus structured observations. Am J Surg. 2015;209:163-169.
19. Molden DC. Understanding priming effects in social psychology: an overview and integration. Soc Cogn. 2014;32:243-249.
20. Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson G. Teaching surgeons about non-technical skills. Surg J R Coll Surg Edinb Irel. 2007;5:86e89.
21. Unknown. Definitive Surgical and Anaesthetic Trauma Care Course.
22. Boffard K. Manual of Definitive Surgical Trauma Care. 1 ed. London: Edward Arnold; 2003.
23. Bleeker CP. Definitive Anaesthetic Trauma Care Courses.
24. Wetzel CM, Kneebone RL, Woloshynowych M, et al. The effects of stress on surgical performance. Am J Surg. 2006;191:5-10.
25. Arora S, Hull L, Sevdalis N, et al. Factors compromising safety in surgery: stressful events in the operating room. Am J Surg. 2010;199:60-65.
26. Anton NE, Montero PN, Howley LD, Brown C, Stefanidis D. What stress coping strategies are surgeons relying upon during surgery? Am J Surg. 2015;210: 846-851.
27. LeBlanc VR. The effects of acute stress on performance: implications for health professions education. Acad Med J Assoc Am Med Coll. 2009;84:S25-eS33.
28. Sadideen H, Kneebone R. Practical skills teaching in contemporary surgical education: how can educational theory be applied to promote effective learning? Am J Surg. 2012;204:396-401.
29. Dedy NJ, Fecso AB, Szasz P, Bonrath EM, Grantcharov TP. Implementation of an effective strategy for teaching nontechnical skills in the operating room: a single-blinded nonrandomized trial. Ann Surg. 2016;263:937-941.
30. Eppich WJ, Hunt EA, Duval-Arnould JM, Siddall VJ, Cheng A. Structuring feedback and debriefing to achieve mastery learning goals. Acad Med J Assoc Am Med Coll. 2015;90:1501-1508.
31. van Merrienboer JJG, Clark RE, de Croock MBM. Blueprints for complex learning: the 4C/ID-model. Educ Technol Res Dev. 2002;50:39-61.
32. Palter VN, Grantcharov TP. Simulation in surgical education. CMAJ Can Med Assoc J. 2010;182:1191-1196.
33. Janssen-Noordman AM, Merrienboer JJ, van der Vleuten CP, Scherpbier AJ. Design of integrated practice for learning professional competences. Med Teach. 2006;28:447-452.
34. Gaba DM. The future vision of simulation in health care. Qual Saf health care. 2004;13(Suppl 1):i2-10.
35. LeBlanc VR, Bould MD, McNaughton N, Brydges R, Piquette D, Sharma B. Simulation in Postgraduate Medical Education. Members of the FMEC PG consortium. 2011.

CHAPTER 5

Stress measurements in surgeons and residents using a smart patch

Mariska Weenk[1], Alexander Alken[1], Lucien Engelen[2], Sebastian Bredie[3], Tom van de Belt[2], Harry van Goor[1]

[1]Department of Surgery
[2]Radboud REshape Innovation Center
[3]Department of Internal Medicine

Radboud University Medical Center, Nijmegen, The Netherlands

The American Journal of Surgery 2018; 216 (2): 361-368

ABSTRACT

BACKGROUND: Stress may negatively affect surgeons' performance during surgical procedures, jeopardizing patient safety. For measuring stress, complex methods are used that cannot record stress real time. This study reports stress measurements in surgeons and residents using a novel patch sensor to identify activities and risk factors of stress.

METHODS: In this explorative study, surgeons and residents wore the HealthPatch™ during all daily activities for 2-3 days. The patch recorded heart rate variability (HRV), and real time stress percentage using a validated algorithm of heart rate (HR) and HRV. The patch was compared with self perceived stress reporting using State Trait Anxiety Inventory (STAI).

RESULTS: A significant increase in HRV and stress percentage was shown in twenty surgeons and residents during surgery in comparison with other activities. Consultants showed lower stress levels while operating compared to fellows and residents. Stress according to the patch did not correlate with STAI outcome.

CONCLUSIONS: Continuous stress monitoring using a wearable sensor patch reveals relevant data on actual stress of surgeons and residents. Stress was highest performing an operation, particularly in fellows and residents.

INTRODUCTION

Surgery is a stressful profession.[1] Long and continuous working hours, high workload, dealing with life and death,[1] and technically challenging procedures[2] are common contributors to stress in surgeons and surgical residents. Chronic stress can lead to relational issues, depression, and burnout,[3-5] and also increases the risk of cardiovascular disease and decreases life expectancy.[6-9]

Surgeons and residents spend a large part of their time in the operating room. Stress can both positively and negatively affect surgeon's performance in the operating room. While moderate levels of stress are necessary to improve alertness, focus, efficiency of action and thus overall performance ('good stress'),[10] excessive and long lasting stress is known to compromise technical[11-13] and non-technical skills ('bad stress').[13] During surgical procedures, excessive levels of stress are mainly caused by technical problems, complexity of the procedure, equipment failures, patient complications, interruptions, and high workload.[10,14,15] During laparoscopic procedures, stress is associated with a prolonged operation time,[11,12] poorer motion efficiency, and an increased number of errors.[12,16] Excessive levels of stress also impair non-technical skills such as communication,[10,17,18] teamwork, judgment, and decision making.[10,17] Loss of these abilities is associated with undesirable events in the operating room and could compromise patient safety.[13-15,18-22]

Research of surgical stress has been focused on the operating room environment and stress has rarely been studied during other activities.[23,24] However, ward rounds and seeing patients in the Emergency Room, the Intensive Care Unit and outpatient clinic may also elicit stress, of which the consequences for the quality of work are unknown.

Heart rate (HR), heart rate variability (HRV), skin conductance, eye blinks, and salivary cortisol[13] are objective markers for stress response. HRV in particular has shown to be a reliable and more time related measure for stress than the other markers.[25] Several studies showed changes in HRV recordings in surgeons during surgical procedures, indicating an increase in intra-operative stress.[24,26-31] Qualitative measurement of stress is commonly by self-reporting questionnaires, such as the State Trait Anxiety Inventory (STAI).[32] Arora et al.[33] developed a method to measure surgeons' stress during surgical procedures using the so-called Imperial Stress Assessment Tool (ISAT). By combining heart rate, salivary cortisol and STAI data, they were able to measure intra-operative stress in a reliable and valid manner. Drawbacks of this tool are the complexity, the time consuming and expensive cortisol analyses, and the inability to obtain real-time stress levels and for a longer period.

Recently, wearable sensors became available for use in healthcare, which can continuously measure vital signs such as HR in an easy and reliable way. The Vital Connect HealthPatch™ (current version is called VitalPatch®, Campbell, CA, USA) is a small, lightweight and comfortable patch, which is attached to the chest. The patch is unique in measuring stress continuously and depicting stress real time, using a validated algorithm that computes stress

as a combination of HRV and HR.[34,35] Because of these features the patch has potential to be used in training situations and to assess chronic stress.
An exploratory study was conducted determining the value of the patch in continuously measuring stress levels in surgeons and surgical residents during all work activities in comparison to usual self perceived stress scoring. Important objective was to evaluate to what extent demographic and surgical factors, surgical experience level in particular, affect this stress.

METHODS

Participants

Consultants, fellows and residents were recruited from the surgical department of the Radboud university medical center in the Netherlands between July 2014 and December 2014. A sample size calculation was not deemed necessary because of the exploratory nature of this study. Demographics including gender, age, level of surgical experience and concurrent use of medication affecting heart rate were noted. Participants gave verbal consent after being informed about the study and the anonymous reporting of data. After reviewing the study protocol, the institutional review board waived the need for formal review and approval (2014-1603).

Patch details

The HealthPatch™ is a flexible self-adhesive patch containing two ECG electrodes and a battery (Figure 1). The patch is validated to measure nine items: single-lead ECG, HR, HRV, stress level in percentage, respiratory rate, skin temperature, body posture, activity and steps. Patch data are streamed to a smart phone via Bluetooth, from where they are transmitted to a secured online cloud for storage. Data can be downloaded from the individual accounts for analysis.

Procedure

Participants wore the patch for at least 48 h. In the morning of the first day, the patch was attached to the participant's chest and a connection was established between the patch and a smart phone via Bluetooth. Baseline patch data and STAI score were collected during 15 min of rest in which participants were instructed to sit comfortably, not performing any physical activity, not reading or speaking. Thereafter, data were continuously collected during all daily work activities for the next 48-72 h. Participants filled out the STAI before and immediately after each surgical procedure, not before and after other activities. This was decided because the other daily activities are more heterogeneous and more frequent e.g. administrative activity. All participants kept a personal logbook in which they documented the type and time of daily activities and also physical activity (e.g. running, taking stairs). At the end of each working day one researcher (MW) debriefed participants with help of the personal logbook. Technical failures and side effects of the HealthPatch were documented.

Figure 1. The Vital Connect HealthPatch™.

Stress measurements

HRV

The smart patch measures HRV, which is defined as the variation in time interval between heart beats (R-R interval). The R-R interval is the time between the peaks of two consecutive QRS complexes as recorded by the 125 Hz ECG.[36] According to the recommendations of the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology,[37] subsequent intervals of five minutes were used for automatic calculation of the HRV by the patch. This was done by using time domain and frequency domain analyses. For time domain, the standard deviation of the interval between two heart beats (SDNN) and square root of the mean R-R interval (RMSSD) were used as parameters reflecting HRV. Low SDNN or RMSSD indicate stress.[38] In the frequency domain, Fourier transformation was used by the patch to calculate the power spectral density. Three main spectral densities were distinguished: the very low frequency (VLF) component (0-0.04 Hz), the low frequency (LF) component (0.04-0.15 Hz) and the high frequency (HF) component (0.15-0.40 Hz). LF and HF represent two branches of the autonomic nervous system; LF is expected to be a marker of sympathetic modulation with some parasympathetic act and HF is a marker of vagal modulation.[39,40] Stress is accompanied by an increase of sympathetic activity, resulting in an increase in LF and a decrease in HF.[38,40] The LF/HF ratio was calculated separately by a researcher to isolate sympathetic tone more precisely.

Real time stress percentage

The patch shows stress real time every four seconds.[34] This stress depicts the result of an algorithm that uses HR and SDNN: Stress (%) = HR + α * SDNN.[34] This stress algorithm was validated and has shown to be sensitive for acute changes in psychological stress.[35] When stress occurs, HR will increase and SDNN will decrease. According to the manufacturer α is usually a negative number. The stress percentage is calibrated to the individual baseline HR. This is done by mapping stress to a cumulative distributive function (Gaussian CDF), ranging between 0 and 1 and multiplied by 100. The stress shown is also adapted to the personal range of daily stress by adjusting the Gaussian CDF to new stress data. The lowest stress level is '0' and highest stress level is '100'. The patch stops measuring the stress percentage when physical activity e.g. walking stairs is undertaken. Thus, only mental stress is recorded by the stress percentage.

Stress perception

For stress perception, the short version of the STAI was used (Table 1).[41] This validated test consists of six items on a four-point scale and measures emotional, cognitive and physical stress. The STAI takes about 2 min to complete. Total STAI scores range from 6 to 24, whereby higher scores indicate an increase in perceived stress.

Table 1. State trait anxiety inventory.

	Not at all	Some-what	Moderately so	Very much
I feel calm	1	2	3	4
I feel tense	1	2	3	4
I feel upset	1	2	3	4
I am relaxed	1	2	3	4
I am content	1	2	3	4
I am worried	1	2	3	4

Data analysis

Participants were divided into groups according to gender and level of experience; consultants (two or more years of independent practice), fellows (surgeons with less than two years of independent practice), senior residents (postgraduate year (PGY) 5 or 6), and junior residents (PGY 4 or less). For analysis fellows and senior residents were grouped together. Daily activities of participants were divided in baseline, surgical procedures, and non-surgical activities (ward visits, outpatient clinic, and administrative work). Time with no clinical activities according to the personal logbook was excluded from further analysis. All surgical procedures performed by the participant during the day or evening were included. Surgical procedures were divided into short (<2 h) and long procedures (2 h) as a proxy for complexity of the operation and hypothetically a difference in stress. Only elective surgical procedures were included. All baseline, outpatient and ward activities were included in the analysis. For administrative work activities one representative period per participant was selected based on the logbook. Data were downloaded in *.CSV files (MS Excel 2007). Raw data were inspected for artifacts and further analyzed in SPSS version 20.0 (SPSS, Inc, Chicago, IL). A surgical procedure was defined as stressful when the postoperative STAI score was at least 1 point higher than the preoperative STAI score.[33]

Statistical analysis

All statistical data analyses were performed using MS Excel and SPSS. Descriptive statistics are presented as mean with standard deviation (SD) or median with range, depending on skewness of data distribution. To test for skewness, the Shapiro-Wilk test was used. Gender and duration of operation were compared using the Independent students' t-test or Mann-Whitney test for non-normally distributed data. Statistical significance between the different activities, between levels of experience and STAI scores was calculated using the ANOVA or Kruskal Wallis test. Stressful surgical procedures according to an increase in STAI scores were compared with non-stressful procedures using the independent students' t-test or Mann-Whitney U test. Pearson correlations were used to test for relationships between delta STAI scores and HRV and stress percentage. A p-value less than 0.05 was considered significant.

RESULTS

Five consultants, seven fellows and senior residents, and eight junior residents (11 men and 9 women) participated. The mean age (±SD) of the consultants, fellows and senior residents, and junior residents were respectively 46.20 (±7.16) years, 35.43 (±4.44) years, 32.25 (±1.83) years. The mean (±SD) years of experience at consultant level was 11.80 (±6.91) years. At the time of data collection, one fellow and one junior resident were pregnant for six weeks. One male fellow used beta-blockers. In all participants, data were collected during baseline and administrative work. Measurements involved 63 elective surgical procedures, 22 long and 41 short procedures. Data of eight participants were collected at the outpatient clinic and data of seven participants when at the surgical ward.

In two participants data was missing due to connection failures. In one participant two hours were missing during a surgical procedure, in the other participant four hours were missing during ward visits and administration. In two other participants, the patch lost complete skin contact after two days resulting in data interruptions on the third day. In two other participants, measurements were stopped after two days because of skin irritation. In 16 participants, baseline STAI scores were collected. In 42 of the 63 surgical procedures, STAI was completely filled in before and after the operation. In 21 procedures participants indicated time shortage completing the STAI preparing the next operation.

Stress measurement outcomes

A 40% decrease in SDNN, a 40% decrease in RMSSD, a 64% increase in the LF/HF ratio and a 300% increase in stress percentage were found during surgery in comparison with baseline, indicating increased stress (Table 2). SDNN and RMSSD were decreased and stress percentage increased during surgery in comparison with non-surgical activities (Figure 2). Stress measurement outcomes of non-operative activities did not differ between each other or from baseline.

Table 2. HR, SDNN, HRV and stress percentage during different daily activities.

	Heart rate (bpm) – median (min-max)	SDNN (ms) – mean ± SD	RMSSD (ms) – median (min-max)	LF/HF ratio – mean ± SD	Stress (%) – mean ± SD
Baseline (20 participants)	70.00 (48.65-84.69)[a]	81.03 ± 27.86[b]	38.97 (22.00-131.70)[c]	3.97 ± 2.50[d]	20.04 ± 13.48[e]
Surgery (63 procedures)	87.69 (62.32-120.38)[a, f, j, o]	51.00 ± 12.08[b, g, k, p]	23.50 (13.14-54.21)[c, h, l, q]	6.18 ± 3.04[d, m]	60.77 ± 24.20[e, i, n, r]
Outpatient clinic (8 participants)	74.30 (61.26-82.97)[f]	80.98 ± 9.14[g]	40.39 (25.25-45.63)[h]	4.31 ± 1.91	28.29 ± 14.31[i]
Ward (7 participants)	75.73 (63.98-83.44)[j]	77.30 ± 20.27[k]	34.94 (28.35-63.82)[l]	3.99 ± 1.20[m]	19.48 ± 8.92[n]
Administrative work (20 participants)	72.64 (57.74-88.07)[o]	74.22 ± 25.39[p]	33.52 (18.61-60.09)[q]	4.72 ± 1.92	29.25 ± 19.88[r]

[a]Baseline versus surgery, $p < 0.001$. [b]Baseline versus surgery, $p < 0.001$. [c]Baseline versus surgery, $p = 0.001$. [d]Baseline versus surgery, $p = 0.010$. [e]Baseline versus surgery, $p = 0.001$. [f]Surgery versus outpatient clinic, $p = 0.005$. [g]Surgery versus outpatient clinic, $p = 0.016$. [h]Surgery versus outpatient clinic, $p = 0.001$. [i]Surgery versus outpatient clinic, $p = 0.001$. [j]Surgery versus ward, $p = 0.007$. [k]Surgery versus ward, $p = 0.006$. [l]Surgery versus ward, $p = 0.006$. [m]Surgery versus ward, $p = 0.408$. [n]Surgery versus ward, $p < 0.001$. [o]Surgery versus administrative work, $p < 0.001$. [p]Surgery versus administrative work, $p < 0.001$. [q]Surgery versus administrative work, $p < 0.001$. [r]Surgery versus administrative work, $p < 0.001$.

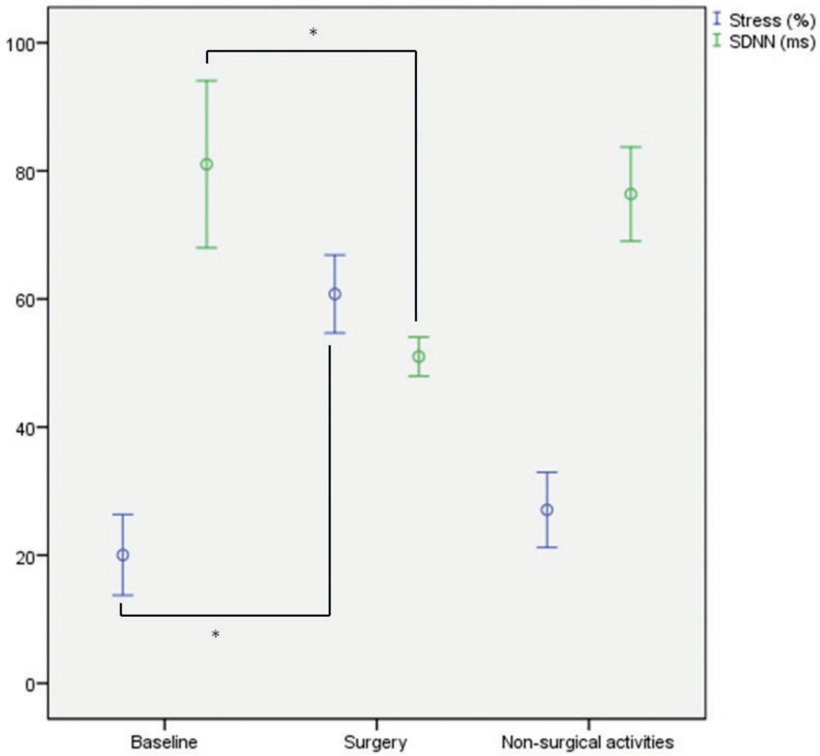


Figure 2 Mean stress percentage and SDNN with 95% confidence intervals between activities. *p < 0.001.

Demographic and surgical factors

Baseline stress measurements outcomes were comparable between men and women (Table 3). SDNN and RMSSD were significantly lower in women compared to men during surgery, also when excluding the two pregnant females and the male using beta-blockers. SDNN, RMSSD, LF/HF ratio, and stress percentage were comparable between long and short surgical procedures (50.58 ± 14.18 vs. 51.23 ± 10.97, $p = 0.250$; 25.94 (14.53-54.21) vs. 23.46 (13.14-51.70), $p = 0.697$; 6.33 ± 3.38 vs. 6.10 ± 2.89, $p = 0.526$; 63.43 ± 23.82 vs. 59.34 ± 24.58, $p = 0.451$, respectively).

During surgery, fellows and senior residents had higher stress percentages and lower SDNN and RMSSD scores than consultants (Table 4; Figure 3). Lower RMSSD scores were also found in junior residents. These results indicate higher stress during surgery by less experienced participants. Three examples of the stress course of a day of operations and of outpatient clinic activities combined with debriefing data are given in Figure 4.

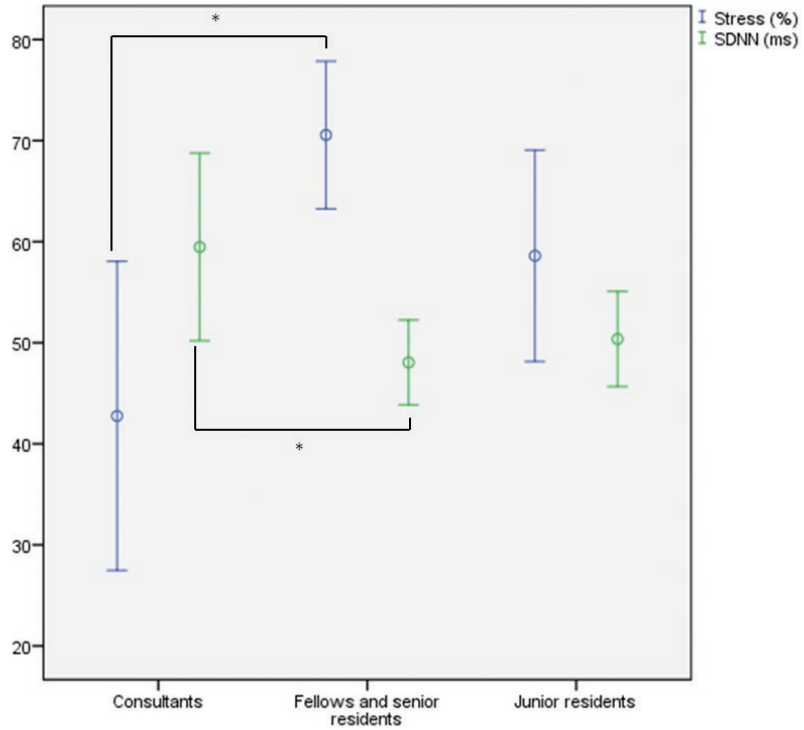


Figure 3 Mean stress percentage and SDNN with 95% confidence intervals between levels of experience. *p < 0.001.

Stress perception

Baseline STAI score was higher in men than in women (9.67 ± 1.66 and 6.70 ± 0.95; p = 0.001). Levels of experience did not affect baseline STAI score. Significant difference was found between baseline STAI scores and preoperative STAI scores (8.38 ± 2.03 vs. 10.12 ± 2.85; p = 0.043). Fifteen of the 42 surgical procedures with complete STAI data were identified as stressful. Gender or level of experience did not differ between perceived stressful and non-stressful procedures. SDNN, RMSSD, LF/HF ratio and stress percentage did not differ between stressful and non-stressful procedures (48.05 ± 7.09 vs. 51.47 ± 11.86, p = 0.250; 25.34 (16.19-37.79) vs. 21.00 (13.14-54.21); p = 0.705; 6.13 ± 3.37 vs. 7.08 ± 3.32, p = 0.386; and 61.00 ± 22.93 vs. 61.98 ± 24.63, p = 0.898, respectively). Delta STAI scores did not correlate with SDNN (r = -0.212, p = 0.178), RMSSD (r = 0.022, p = 0.892), LF/HF ratio (r= 0.033, p = 0.835) and stress percentage (r = -0.046, p=0.771).

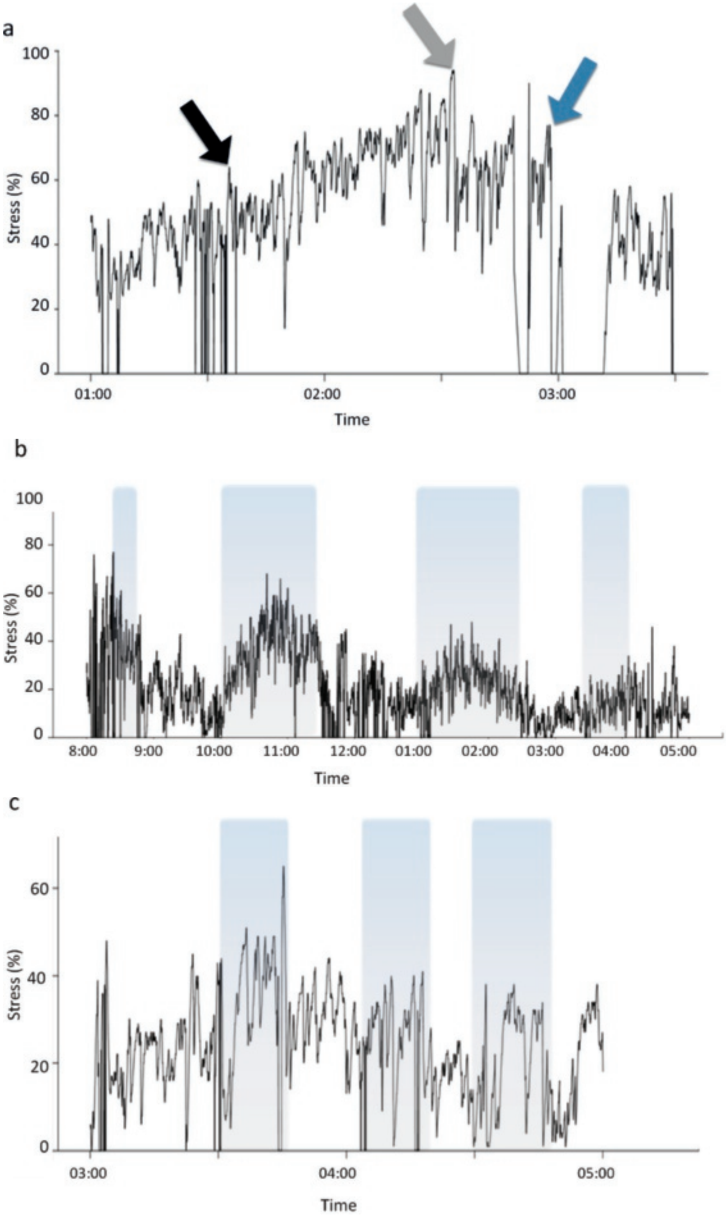


Figure 4 a: Stress pattern (%) of a fellow during a surgical procedure (hemihepatectomy). Black arrow indicates start of operation; grey arrow indicates when senior surgeon enters the OR for supervision; blue arrow indicates end of the hardest part of the operation. **b:** Stress pattern (%) of a consultant during four surgical procedures (blue areas). **c:** Stress pattern (%) of a consultant during an afternoon of outpatient clinic. Blue areas indicate administrative work in between patient contact using a newly introduced electronic medical record system. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article)

Table 3. HR, HRV and stress percentage between men and women

	Baseline		Surgery	
	Men	**Women**	**Men**	**Women**
Heart rate (bpm) – median (min-max)	68.57 (48.65-84.69)	70.61 (59.50-75.46)	86.18 (62.32-106.42)[a]	94.73 (68.98-120.38)[a]
SDNN (ms) – mean ± SD	87.94 ± 30.31	72.57 ± 23.41	54.69 ± 11.66[b]	46.67 ±11.26[b]
RMSSD (ms) – median (min-max)	47.98 (22.00-131.70)	31.29 (22.57-77.78)	27.75 (16.19-54.21)[c]	21.00 (13.14-42.69)[c]
LF/HF ratio – mean ± SD	3.95 ± 2.76	4.00 ± 2.31	6.22 ± 3.16	6.13 ± 2.96
Stress (%) – mean ± SD	16.63 ± 12.25	24.19 ± 14.45	55.35 ± 20.56	67.12 ± 26.86

SDNN = standard deviation of the R-R intervals; RMSSD = square root of the mean R-R interval; LF = low frequency; HF = high frequency. ap=0.050. bp=0.007. cp=0.001.

DISCUSSION

Continuous stress monitoring in surgeons and surgical residents using a new, small and light weighted, wearable sensor patch reveals relevant and almost complete data on stress levels of surgeons and residents during their workday. Performing an operation was more stressful than other daily activities, particularly for fellows and residents, and based on different stress calculations. The patch did not interfere with the activities and was well tolerated by most participants, and measured actual time-related stress levels differences during real life situations, whereas a common subjective stress evaluation (STAI) did not find any difference.

We used various calculations for determining stress levels based on heart rate and heart rate variability data and the real time depicted stress percentage. Stress percentage results compared well with the calculated data. This outcome favors the use of stress percentage level as indicator of stress with this patch because this parameter is real time depicted, easy to read, can show rapid changes and is independent of physical activity. For the first time continuous (self)monitoring of stress for a long period is possible which could not be achieved with existing methods such as the Imperial Stress Assessment Tool and saliva cortisol.[33] Parameters for parasympathetic and sympathetic activity (LF and HF) showed less significant differences and only between operations and baseline values. LF/HF ratio is more sensitive for artifacts and is less reliable over shorter periods.[42] However, results should be interpreted with caution due to low numbers in subgroup analyses and possible inaccuracy in LF and HF data with a relatively low sample frequency of the patch.

Table 4. Recordings during surgical procedures; divided by level of experience

	Consultants (n=5)	Fellows and senior residents (n=7)	Junior residents (n=8)
Heart rate (bpm) – median (min-max)	75.00 (62.32-97.17)[a, e]	93.53 (71.18-120.38)[a]	88.98 (71.54-107.10)[e]
SDNN (ms) – mean ± SD	59.47 ± 13.82[b]	48.05 ± 10.38[b]	50.36 ± 11.68
RMSSD (ms) – median (min-max)	31.55 (22.75-38.27)[c, f]	20.55 (13.14-54.21)[c]	22.54 (13.69-51.70)[f]
LF/HF ratio – mean ± SD	5.60 ± 1.61	6.59 ± 3.41	6.02 ± 3.16
Stress (%) – mean ± SD	42.76 ± 22.76[d]	70.56 ± 18.09[d]	58.60 ± 25.90

SDNN = standard deviation of the R-R intervals; RMSSD = square root of the mean R-R interval; LF = low frequency; HF = high frequency. [a]Consultants versus Fellows and senior residents, p=0.001. [b]Consultants versus Fellows and senior residents, p=0.024. [c]Consultants versus Fellows and senior residents, p=0.018. [d]Consultants versus Fellows and senior residents, p=0.003. [e]Consultants versus Junior residents, p=0.002. [f]Consultants versus Junior residents, p=0.036.

High stress levels during operations have been reported[24,26-30] but comparisons with other daily activities was not yet examined. Performing an operation gave more mental stress than activities at the ward and the outpatient clinic, or when doing administration. Interpretation of these differences should be done with caution. Less than half of the participants had outpatient clinic or ward activities during the days the patch was worn and there could have been a selection bias in other activities. Also one period of administrative activity per person was taken into account albeit representative for these activities. Previous studies also showed lower stress levels in experienced surgeons in comparison with younger colleagues.[13,29,43-48] This is possibly explained by differences in coping strategies.[10] Consultants seem more capable of recognizing internal signals indicating stress, such as heart pounding and clouded judgment, and may have developed better coping strategies to deal with stress by for example physical relaxation methods, distancing techniques and self-talk. In contrast to senior residents and fellows, no significant difference in stress levels was found between consultants and junior residents. This could possibly be explained by the fact that junior residents operate under more supervision than senior residents and fellows.

The increased stress during operation may be considered as 'good stress', reflecting increased focus. We cannot rule out that some elevations of the stress are due to increased focus and excitement. Due to the small sample size we could not analyze stress data in relation to intra-operative problems, which might indicate 'bad stress'. However, we observed long lasting stress levels during surgery corresponding with increased stress levels during other, potentially, more relaxing activities in some individuals, which may indicate 'bad (chronic) stress'.

It was not an aim of this study to validate objective stress measured by this device against a subjective self-perceived stress by the STAI questionnaire. We, however, compared outcomes of these two different stress measurements to have an impression of their relationship. As shown in the results section no correlation was found between subjective and objective stress measurements. Underreporting of perceived stress in general and specifically in surgical specialists has been reported.[49] Perceived stress might have been affected by a short moment of the procedure and is dependent on recollection after the operation, whereas objective stress calculations encompassed the total operation and were expressed in mean or median. In contrast to STAI the device is potentially more suitable to pick up more and longer 'unnoticed' stress moments, which is relevant for determination of chronic stress.

Strength of this study is the comparison of stress obtained by continuous registration, between all daily activities in a group of surgeons and residents and during two to three days in a row. Combining stress data with notes in the logbooks and daily debriefings also allowed for detailed insight in individual stress patterns. Frequent and long during high stress percentages were found in some individuals and at more than one daily activity (see Figure 4).

The small number of participants, the few demographics obtained and the missing patch and STAI data limited further subgroup analyses for stress risk factors. Measurements prematurely stopped in 20% of the participants due to patch dysfunction or irritation. The skin irritation in 10% of the participants would hamper use of the patch for more days in a row. Adhesive patches for sensitive skin are developed and may decrease skin irritation. Future studies should take the limitation of occasional patch dysfunction into account.

Regarding its ability to measure stress continuously and depict stress real time this sensor device has large potential in healthcare, both for healthcare workers and patients, both in daily practice and in a training environment, and both for an individual and a team effort. Particularly trainees may benefit recognizing stressors and stressful situations real time and learning to cope with or prevent stress. Operating room team simulation training using the patch in all participants would allow residents and consultants to train various crisis situations, to experience stress and to reflect on the consequences regarding quality and safety of the operation. Other potential application is the early and simple assessment of chronic stress in patients and healthcare workers by computer analysis of continuous or serial time periods of patch data. Other means for chronic stress analysis such as hair analysis are costly and still need validation.[50,51] Ongoing studies focus on stress monitoring in trainees and faculty during surgical simulation training and in patients and nurses at the surgical ward.

CONCLUSION

Continuous stress monitoring in surgeons and residents using a simple, easy to wear sensor patch reveals real time data on different stress levels of surgeons and residents during the day. With this new technique we could demonstrate that performing an operation is more stressful than other daily activities in the hospital, particularly with less surgical experience. The stress percentage allows for real time feedback of the stress level making the sensor patch suitable for a widespread application in healthcare.

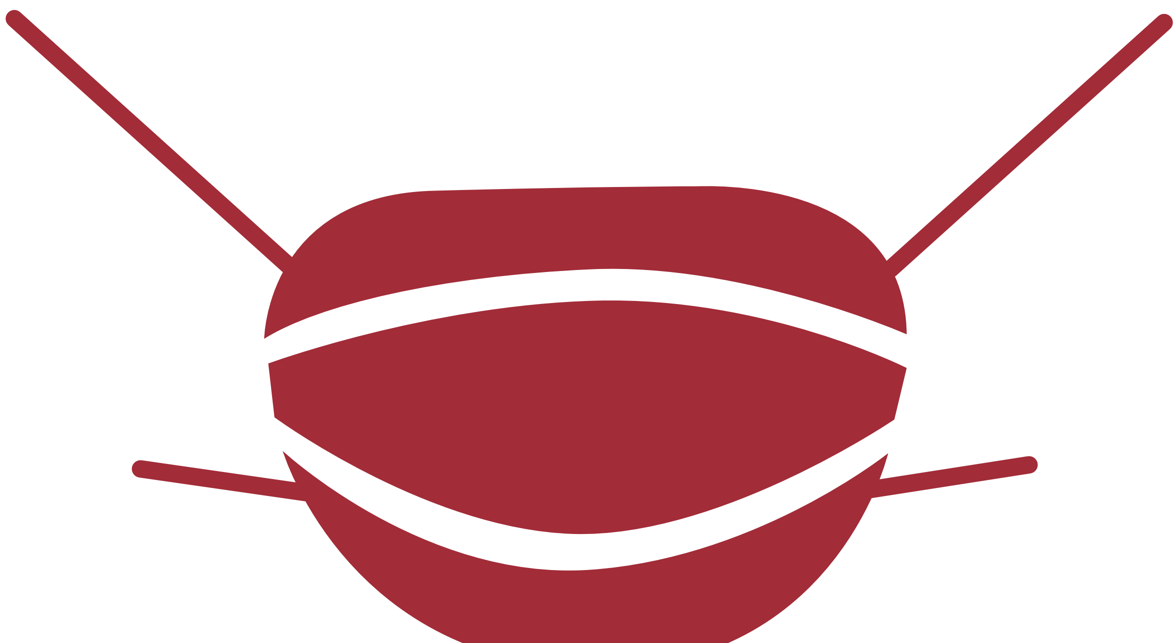
REFERENCES

1. Balch CM, Freischlag JA, Shanafelt TD. Stress and burnout among surgeons: understanding and managing the syndrome and avoiding the adverse consequences. Arch Surg 2009;144(4):371-376.
2. Khajuria A, Maruthappu M, Nagendran M, et al. What about the surgeon? Int J Surg 2013;11(1):18-21.
3. Oskrochi Y, Maruthappu M, Henriksson M, et al. Beyond the body: A systematic review of the nonphysical effects of a surgical career. Surgery 2016;159:650-664.
4. Campbell Jr DA, Sonnad SS, Eckhauser FE, et al. Burnout among American surgeons. Surgery 2001;130(4):696-705.
5. Taylor C, Graham J, Potts HWW, et al. Changes in mental health of UK hospital consultants since the mid-1990s. Lancet 2005;366(9487):742-744.
6. Steptoe A, Kivimaki M. Stress and cardiovascular disease: an update on current knowledge. Annu Rev Publ Health 2013;34:337-354.
7. Brotman DJ, Golden SH, Wittstein IS. The cardiovascular toll of stress. Lancet 2007;370(9592):1089-1100.
8. Eller NH, Netterstrom B, Gyntelberg F, et al. Work-related psychosocial factors and the development of ischemic heart disease: a systematic review. Cardiol Rev 2009;17(2):83-97.
9. Richardson S, Shaffer JA, Falzon L, et al. Meta-analysis of perceived stress and its association with incident coronary heart disease. Am J Cardiol 2012;110(12):1711-1716.
10. Wetzel CM, Kneebone RL, Woloshynowych M, et al. The effects of stress on surgical performance. Am Journ Surg 2006;191:5-10.
11. Hassan I, Weyers P, Maschuw K, et al. Negative stress-coping strategies among novices in surgery correlate with poor virtual laparoscopic performance. Br J Surg 2006;93(12):1554-1559.
12. Arora S, Sevdalis N, Aggarwal R, et al. Stress impairs psychomotor performance in novice laparoscopic surgeons. Surg Endosc 2010;24(10):2588-2593.
13. Arora S, Sevdalis N, Nestel D, et al. The impact of stress on surgical performance: A systematic review of the literature. Surgery 2010;147(3):318-330.
14. Arora S, Hull L, Sevdalis N, et al. Factors compromising safety in surgery: stressful events in the operating room. Am J Surg 2010;199(1):60-65.
15. Anton NE, Montero PN, Howley LD, et al. What stress coping strategies are surgeons relying upon during surgery? Am J Surg 2015;210(5):846-851.
16. Moorthy K, Munz Y, Dosis A, et al. The effect of stress-inducing conditions on the performance of a laparoscopic task. Surg Endosc 2003;17(9):1481-1484.
17. Arora S, Sevdalis N, Nestel D, et al. Managing intraoperative stress: what do surgeons want from a crisis training program? Am J Surg 2009;197(4):537-543.
18. Lingard L, Espin S, Whyte S, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. Qual Saf Health Care 2004;13(5):330-334.

19. Christian CK, Gustafson ML, Roth EM, et al. A prospective study of patient safety in the operating room. Surgery 2006;139(2):159-173.
20. Rogers SO, Jr., Gawande AA, Kwaan M, et al. Analysis of surgical errors in closed malpractice claims at 4 liability insurers. Surgery 2006;140(1):25-33.
21. Griffen FD, Stephens LS, Alexander JB, et al. Violations of behavioral practices revealed in closed claims reviews. Ann Surg 2008;248(3):468-474.
22. Gawande AA, Zinner MJ, Studdert DM, et al. Analysis of errors reported by surgeons at three teaching hospitals. Surgery 2003;133(6):614-621.
23. Becker WGE, Ellis H, Goldsmith R, et al. Heart rates of surgeons in theatre. Ergonomics 1983;26(8):803-807.
24. Demirtas Y, Tulmac M, Yavuzer R, et al. Plastic Surgeon's Life: Marvelous for Mind, Exhausting for Body. Plast Reconstr Surg 2004;114(4):923-931.
25. Chandola T, Heraclides A, Kumari M. Psychophysiological biomarkers of workplace stressors. Neurosci Biobehav R 2010;35(1):51-57.
26. Jones KI, Amawi F, Bhalla A, et al. Assessing surgeon stress when operating using Heart Rate Variability and the State Trait Anxiety Inventory: will surgery be the death of us? Colorectal Dis 2014;17(4):335-341.
27. Prichard RS, O'Neill CJ, Oucharek JJ, et al. A prospective study of heart rate variability in endocrine surgery: surgical training increases consultant's mental strain. J Surg Educ 2012;69(4):453-458.
28. Song MH, Tokuda Y, Nakayama T, et al. Intraoperative heart rate variability of a cardiac surgeon himself in coronary artery bypass grafting surgery. Interact Cardiovasc Thorac Surg 2009;8(6):639-641.
29. Böhm B, Rötting N, Schwenk W, et al. A prospective randomized trial on heart rate variability of the surgical team during laparoscopic and conventional sigmoid resection. Arch Surg 2001;136(3):305-310.
30. Czyżewska E, Kiczka K, Czarnecki A, et al. The surgeon's mental load during decision making at various stages of operations. Europ J Appl Physiol 1983;51(3):441-446.
31. Rieger A, Stoll R, Kreuzfeld S, et al. Heart rate and heart rate variability as indirect markers of surgeons' intraoperative stress. Int Arch Occ Env Hea 2014;87(2):165-174.
32. Speilberger CD, Gorsuch RL, Lushene R. STAI manual. Palo Alto, CA: Consulting Psychologist Press; 1970.
33. Arora S, Tierney T, Sevdalis N, et al. The Imperial Stress Assessment Tool (ISAT): a feasible, reliable and valid approach to measuring stress in the operating room. World J Surg 2010;34(8):1756-1763.
34. Vital Connect [personal communication]. Vital Connect Stress Algorithm. 2014
35. Selvaraj N. Psychological acute stress measurement using a wireless adhesive biosensor. Conf Proc IEEE Eng Med Biol Soc 2015;2015:3137-3140.
36. Selvadaj N. Long-term Remote Monitoring of Vital Signs using a Wireless Patch Sensor. Health Innovations and Point-of-Care Technologies Conference. 2014.
37. Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology. Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use. Circulation 1996;93(5):1043-1065.

38. Delaney JPA, Brodie DA. Effects of short-term psychological stress on the time and frequency domains of heart-rate variability. Percept Motor Skill 2000;91(2):515-524.
39. Routledge FS, Campbell TS, McFetridge-Durdle JA, et al. Improvements in heart rate variability with exercise therapy. Can J Cardiol 2010;26(6):303-312.
40. Malliani A, Pagani M, Lombardi F, et al. Cardiovascular neural regulation explored in the frequency domain. Circulation 1991;84(2):482-492.
41. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). Brit J Clin Psychol 1992;31(3):301-306.
42. Allen J. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. IEEE Trans Acoust 1977;25(3):235-238.
43. Thielmann B, Boeckelmann I. Heart Rate Variability as an Indicator of Mental Stress in Surgeons - A Review of the Literature. Zentralbl Chir 2016;141(5):577-582.
44. Klein M, Mouraviev V, Craig C, et al. Mental stress experienced by first-year residents and expert surgeons with robotic and laparoscopic surgery interfaces. J Robotic Surg 2014;8(2):149-155.
45. Kikuchi K, Okuyama K, Yamamoto A, et al. Intraoperative stress for surgeons and assistants. J Ophtalmic Nurs Technol 1995;14(2):68-70.
46. Yamamoto A, Hara T, Kikuchi K, et al. Intraoperative stress experienced by surgeons and assistants. Ophthalmic surg lasers 1999;30(1):27-30.
47. Berguer R, Smith WD, Chung YH. Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. Surg Endosc 2001;15(10):1204-1207.
48. Marrelli M, Gentile S, Palmieri F, et al. Correlation between Surgeon's experience, surgery complexity and the alteration of stress related physiological parameters. PloS one 2014;9(11):e112444.
49. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. BMJ 2000;320(7237):745-749.
50. Russell E, Koren G, Rieder M, et al. Hair cortisol as a biological marker of chronic stress: current status, future directions and unanswered questions. Psychoneuroendocrinology 2012;37(5):589-601.
51. Iglesias S, Jacobsen D, Gonzalez D, et al. Hair cortisol: A new tool for evaluating stress in programs of stress management. Life Sci 2015;141:188-192.

5



CHAPTER 6

Measuring stress and coaching behaviors during a highly realistic trauma surgery team training

Alexander Alken[1], Cornelia Fluit[2], Mariska Weenk[1], Mats Koeneman[3], Jan-Maarten Luursema[1], Harry van Goor[1]

[1]Department of Surgery
[2]Radboudumc Health Academy
[3]Radboud REshape Innovation Center and

Radboud University Medical Center, Nijmegen, The Netherlands

Submitted

ABSTRACT

BACKGROUND: The operating room (OR) is a dynamic, high-stakes teaching environment. Excessive stress in surgical teachers can result in disruptive coaching which may harm trainees' learning. Little is known about the impact of stress on teaching behavior. We analyzed differences in coaching behavior during episodes of high and low stress in teachers during a highly realistic trauma surgery team training. We also assessed the minute-to-minute relationship between salient training events and stress levels of a teacher and a trainee.

METHODS: Surgical trainees trained six emergency surgery procedures on anesthetized porcine models while working in operating teams. In each team one surgical teacher coached two trainees. Nine experienced teachers and one trainee were equipped with an innovative wearable sensor measuring Low Frequency / High Frequency ratio on ECG recordings as a proxy for stress. All coaching activities were recorded on audio and video. For each teacher, four 10-minute time frames were determined representing the two lowest and the two highest mean Low Frequency / High Frequency ratios. A blinded observer categorized and calculated the frequency of all technical and non-technical coaching behaviors using a previously developed observational instrument. To explore the relationship between salient events and stress, two 30-minute time periods were selected with steep increases and decreases in the Low Frequency / High Frequency ratio of one teacher and one trainee working in the same team, and compared with the occurrences of salient events collected by a blind observer.

RESULTS: Useable data was collected in five teachers and one trainee. We found no differences in the teachers' frequency of coaching behaviors during their highest and lowest Low Frequency / High Frequency ratio. For the trainee, steep increases in the Low Frequency / High Frequency ratio coincided with the infliction of injuries. No such clear pattern emerged for the teacher.

CONCLUSIONS: Stress, as measured by an objective and unobtrusive wearable device, was not related to coaching behaviors of experienced teachers regarding the attention to non-technical skills in a highly realistic trauma surgery team training. Trainee stress appeared to be time related to the infliction of injuries. Our method combining objective stress measurements with observations may stimulate research relating stress to teachers' and trainees' behaviors, performance and stress coping.

INTRODUCTION

The operating room (OR) is a dynamic, high-stakes and stressful environment. At the same time, it is only in the OR that surgical trainees can develop the skills, knowledge and strategies to cope with and work in such a complex environment. Coaching by experienced surgeons is an essential prerequisite for the OR as a training environment.[1] The OR holds many potential stressors which can easily compromise the non-technical team and communication skills of a surgeon and trainee,[2-6] and probably also teaching. Surgeons under stress can become short-tempered[6] and use disruptive communication styles,[3,7] which have been reported to hinder learning.[8] A recent study demonstrated that teachers applying a harshly criticizing and belittling communication style jeopardize the trainees' development of laparoscopic skills during simulation training.[9] Trainees, teachers and teaching experts consider a calm, respectful and supportive communication style important for effective teaching.[10,11]

In the current study, we aim to analyze how different levels of objectively measured stress in surgical teachers relate to their coaching behaviors during a highly realistic trauma surgery team training. Teachers' coaching behaviors are often analyzed through subjective trainee or teacher perception questionnaires.[12] These perceptions, however, have been shown to correlate poorly to independent observations by trained observers[12,13] and do not provide insight in what coaching behaviors are exactly applied at what moment in time or in relation to, for example, high or low stress events. Some studies explored a more objective approach and analyzed coaching by categorizing teachers' behaviors with observational instruments.[13-19] Although these observational instruments provide insight in how surgical teachers coach their trainees, all of them have different strong points and weaknesses regarding the amount of distinguished behaviors, the detail of the definitions of behaviors (e.g. questioning as one category or divided into sub categories), and the distinction of skills on which the coaching is focused (e.g. operating techniques or anatomy). An instrument that provides a highly detailed insight in what coaching behaviors are applied when teachers coach their trainees on different surgical skills is still lacking.

Stress in surgeons is often analyzed through subjective perception questionnaires and averaged objective physiological markers, taken after the activity.[4,20] In addition to being collected afterwards, stress perceptions have been shown to be unreliable.[3,7,20,21] A physiological marker, like Heart Rate Variability (HRV; the variability in time between consecutive heartbeats), is objective and has been shown to be a reliable parameter for stress.[22] In previous research, HRV has been used to examine average stress levels over large time periods which limit the insight in how stress relates to specific events. Novel in the current study is the Vital Connect HealthPatch™ (current version is called VitalPatch®, Campbell, CA, USA) which enables the measurement of the surgical teachers' HRV in relation to time.[20]

In the current study, coaching behaviors of surgical teachers are observed and categorized with an instrument based on a previously developed observational instrument.[13,14] This instrument enables a detailed insight into teachers' coaching behaviors. Consequently, the observed coaching behaviors are linked to the stress measurements in teachers in real-time to explore how objectively measured moments of highest and lowest stress are related to the observed coaching behaviors. Based on previous findings of our research group we hypothesize that surgical teachers will coach less frequently on non-technical skills when they are stressed.[23] Since surgeons tend to apply a more assertive and commanding communication style when they experience stress,[6] we also hypothesize that surgical teachers provide less feedback and less detailed explanations when they are experiencing stress. Because stress in surgeons seems to be related to events in the OR such as patient and equipment problems[2] and the surgeon's experience,[20] we also investigate the relationship between events in the OR and the stress experienced by a trainee and a teacher working in one operating team. We expect that the HealthPatch™ enables us to measure HRV as a proxy for stress on a higher temporal resolution than previous physiological recording approaches. This approach should make it possible to relate events and coaching behaviors to stress on an unprecedented temporal resolution.

METHODS

Study design

The study was performed during a hands-on simulation training as part of a trauma surgery course. Surgical teachers were continuously monitored for stress by wearing a Vital Connect HealthPatch™ (current version is called VitalPatch®, Campbell, CA, USA) while they coached trainees who performed various trauma operations. Stress levels of teachers were analyzed and their highest and lowest moments of stress were determined. All training activities were video and audio recorded. The coaching behaviors of teachers were blindly observed during the highest and lowest moments of stress and compared.

Also, one trainee was provided with a HealthPatch™ and was continuously monitored for stress. That way a combined stress graph was depicted for one teacher-trainee duo. Two time periods with several steep increases and decreases of stress were blindly observed for salient events to explore the relation between stress and events in the OR. Further details on the data collection and analyses are described in the following sections.

Course description

The Definitive Surgical and Anesthetic Trauma Care (DSATC) course is a three-day course which combines the Definitive Surgical Trauma Care (DSTC)[24] and Definitive Anesthetic Trauma Care (DATC) course[25]. The course focuses on the training of surgical technical and non-technical team skills.[25,26] Every day starts with 3.5 hours of lectures in the morning followed by a 3.5-hour hands-on workshop in the afternoon. During three days, teachers and trainees rotate over three parallel workshops. One workshop is the integrated acute surgery team training, which is the focus of the current study. Each day, four surgical teachers and eight surgical trainees participate in four OR's in the university's animal research facility encompassing twelve different teachers and 24 different trainees in three days. In each 3.5-hour OR workshop, six acute trauma scenarios are performed on live porcine models. Each scenario consists of one or two complex and potentially fatal injuries inflicted by the surgical teacher. Every surgical teacher coaches two surgical trainees in gaining control of these injuries while they are functioning in an operating team with a scrub nurse, an anesthesiology teacher and one or two anesthesiology trainees. The ethical animal review board approved the DSATC training.

Participants

Nine out of twelve surgical teachers, and 18 surgical trainees, participated in this study. All gave their written informed consent voluntarily. All data was stored in a secured database. Data analysis occurred anonymously. The hospital review board waived the need for formal ethical approval.

Demographics of teachers and trainees

We collected demographics of surgical teachers and trainees including age, use of medication potentially affecting heart rate, PGY-level or years of experience as a surgeon (trainees), years of experience as a surgical teacher and number of completed teach the teacher courses (teachers).

Heart Rate Variability data of teachers

Vital Connect HealthPatch™ devices were attached to the chest at the heart area of the surgical teachers (and one surgical trainee) who participated in the integrated acute surgery team training. The HealthPatch™ is a validated light-weight patch which continuously measures parameters like respiratory rate, heart rate and HRV.[20] All data is transmitted to a secured online data cloud. Only data collected during the integrated acute surgery team training were analyzed.

The HealthPatch™ automatically computes and registers the time passed in milliseconds between consecutive heart beats (the RR-intervals). Based on the collected RR-intervals the Low Frequency / High Frequency ratio (LF/HF-ratio) is calculated, which represents the ratio between the sympathetic and parasympathetic activity (LF) and the vagal modulation (HF).[20] The LF/HF-ratio is a common measure for HRV and a reliable indicator for stress.[22] An increase in LF/HF-ratio means an increase in acute stress. Periods of at least six minutes of consecutive RR-interval recordings without artifacts (failed data registration) are needed to calculate the LF/HF-ratio accurately. One researcher (MK) calculated the mean LF/HF-ratio per teacher over time frames of ten consecutive minutes. Time frames were shifted every minute resulting in mean LF/HF-ratios from minute 0 to 10, 1 to 11, 2 to 12, etcetera. Time frames containing artifacts were discarded. For every teacher MK selected the two most stressful 10-minute time frames (highest and second highest mean LF/HF-ratio) and the two least stressful 10-minute time frames (lowest and second lowest mean LF/HF-ratio).

Coaching behavior of teachers

Teachers' coaching behaviors were recorded on audio via a voice-recorder worn by each teacher and on video via an overview camera in the OR. For every teacher, researcher MK communicated the start and finish times of the two most and two least stressful 10-minute time frames to a second researcher (MW) who was blinded to any stress levels. Researcher MW reviewed the recordings and checked for possible abnormalities (e.g. long periods without a teacher being present). If abnormalities were identified, researcher MK was requested to provide the start and finish times of the second most or least stressful 10-minute time frames, and so forth. Ultimately, the start and finish times of the checked time frames were communicated to a third researcher (AA), a trained observer, who was kept blind for any stress levels. Researcher AA reviewed the checked recordings and scored all teaching behaviors within the time frames.

The coaching behaviors of teachers were scored with an adapted version of a previously reported observational coaching instrument (Table 1).[13,14] This instrument was developed to comprehensively capture the amount of coaching, the skills on which the coaching was directed (A-F), the specificity (I-VIII), and the information density (1-4). Each coaching remark made by a teacher directed to one or both surgical trainees was scored. First, the main category was determined: (A) '*Technical skills*', (B) '*Non-technical skills*', (C) '*Medical and anatomical knowledge*', (D) '*Patient information*', (E) '*Other but relevant*' or (F) '*Irrelevant or indefinable*'. Second, each remark was scored as being (I) '*Instruction*', (II) '*Corrective feedback*', (III) '*Complementing feedback*', (IV) '*Information*', (V) a '*Verifying question*', (VI) a '*Thinking question*', (VII) '*Chitchatting*' or (VIII) '*Other*'. Third, each remark was scored as containing (1) '*No additional information*', (2) basic information on '*How, what or where*', (3) more detailed information about '*Why or when*' or (4) '*Who*'; or a combination of these options. For example, the remark "Be careful, your suturing is too rough. This tissue is very fragile and can easily rip. Don't make your knots too tight." would be coded as *A.II.2.3*. One remark can be scored in more than one category (see example). If a remark was inaudible it was scored as '*Irrelevant or indefinable*'.

Table 1. Observational instrument categories of coaching behavior

Step 1	**Main categories**		
	A	**Coaching directed on trainee's surgical technical skills, actions or future actions**	*- e.g. cutting, suturing, determining surgical approach, etc.*
	B	**Coaching directed on trainee's non-technical communicative and collaborative skills, actions or future actions**	*- e.g. discussing plans with anesthesia, communication with the OR nurse, assisting other surgeon, etc.*
	C	**Coaching directed on trainee's medical or anatomical knowledge**	*- e.g. organ identification, anatomical differences in a pig, working mechanism of haemostasis, etc.*
	D	**Coaching regarding patient information**	*- e.g. patient condition, vital functions, allergies, etc.*
	E	**Other but relevant**	*- e.g. information regarding the progress of the course, etc.*
	F	**Irrelevant or indefinable**	*- e.g. listening responses, inaudible remarks, etc.*
Step 2	**Goal categories**		
	I	**Instruction**	directing a trainee what action he or she has to execute, immediately or in the near future. *- e.g. "First check the small bowel for bleedings"*
	II	**Corrective feedback**	informing a trainee that his or her action was wrong after this action was executed. *- e.g. "You have to do that more carefully."*
	III	**Complementing feedback**	informing a trainee that his or her action was performed well after this action was executed. *- e.g. "Well done! You stopped the bleeding quickly."*
	IV	**Information**	providing information which is not instruction or feedback, e.g. discussing multiple approaches, providing background information, summarizing the case, etc. *- e.g. "If you are with two you can also agree on who is responsible for the upper and the lower part of the intestines according to someone's position towards the patient."*
	V	**Verifying question**	asking a question to check whether a trainee understood the provided information, whether a trainee is still on track, etc. *- e.g. "So, do you know what to do now?"*
	VI	**Thinking question**	asking a question which involves a trainee's knowledge or reasoning. *- e.g. "What do you do if the bleeding continues?"*
	VII	**Chitchatting**	remarks not directly related to the training but relevant for the atmosphere, use of humor, etc. *- e.g. "This pig tried to rob a bank and got shot."*
	VII	**Other**	surplus category
Step 3	**Content categories**		
	1	**No additional information**	remarks are not clarified *- e.g. "That's not okay."*
	2	**How, what, where**	an explanation how something needs to be done, what needs to be done, the location of an organ, etc. *- e.g. "Start with packing."*
	3	**Why, when**	an explanation why or when something needs to be done. *- e.g. "You need to fold the packing materials there because it's too narrow."*
	4	**Who**	an explanation when something needs to be done. *- e.g. "You have to hold the organ aside."*

6

LF/HF-ratio graph of the teacher-trainee duo

An exploratory observation was conducted for one team's surgical teacher-trainee duo (team 3, Table 2). A researcher who was not involved in this study selected two 30-minute time periods during which the teacher's and trainee's LF/HF-ratio values depicted in a graph showed several steep increases and decreases of stress. The start and finish times of both periods were provided to a trained observer (AA) to enable the blind observation of salient events. AA reviewed the recordings and made field notes regarding, for example, the death of the animal, activities like the infliction of injuries by the trainer or cardio pulmonary resuscitation, interactions like silence and tone of voice (for example shouting). Field notes on salient events were qualitatively collected in time frames of one minute and then synchronized in time with the teacher's and trainee's LF/HF-ratio graph.

Data analysis

HRV data of four teachers was lost because the patch lost skin contact frequently while operating (2) or WIFI connection breaches (2). In five surgical teachers sufficient HRV data was collected for analysis. In one of the five teachers 67 percent of the data was lost, in two teachers 26 percent and in the other two teachers 6 percent. In total, 10 highest and 10 lowest moments of stress were analyzed for each of the five teachers' HRV and coaching behavior. Two new time frames had to be selected after initial check due to a lack of teacher presence. Collection of HRV data was successful in the one trainee.

To investigate whether the coaching behaviors of teachers differed during periods of highest and lowest stress, for each surgical teacher, we totaled the coaching behaviors within the two highest stress time frames ('*Total coaching*' high stress condition) and the two lowest stress time frames ('*Total coaching*' low stress condition). Consequently, within each stress condition, the coaching behaviors coded within the main categories A to E were summed up to represent the amount of '*Total relevant coaching*' and expressed in a percentage of the amount of '*Total coaching*' (A to F; Table 1 and 2). Coaching behavior dedicated to non-technical skills (codes B), feedback (codes II and III combined), and highly detailed coaching (all codes with a 3 and/or a 4) were calculated as the '*Total amount of relevant coaching*'. For each teacher, the coaching behavior categories in the highest stress condition were compared to the lowest stress condition. The differences of these variables in both conditions were normally distributed. Statistical differences were analyzed with paired sample t-tests.

To explore the relation between the stress of the teacher and the trainee in relation to salient events in the OR during the training, the field notes were synchronized in time per minute with the LF/HF-ratio graph of the teacher and trainee. We qualitatively analyzed whether increases and decreases in stress matched with the occurrence of salient events.

Table 2. Surgical teachers' HRV data in LF/HF-ratio in relation to the applied coaching behaviors

	Mean LF/HF-ratio (SD)	Highest and lowest LF/HF-ratio		Δ LF/HF-ratio	Total coaching (#)	Total relevant coaching as a percentage of total coaching (#/%)	NTS coaching as a percentage of total relevant coaching (#/%)	Feedback as a percentage of total relevant coaching (#/%)	Highly detailed coaching as a percentage of total relevant coaching (#/%)
1	2.92	**Highest:**	3.60	1.10	101 remarks	73 remarks (72%)	4 remarks (5%)	7 remarks (10%)	11 remarks (15%)
	(0.29)	**Lowest:**	2.50		83 remarks	74 remarks (89%)	2 remarks (3%)	6 remarks (8%)	14 remarks (19%)
2	1.42	**Highest:**	1.68	0.55	76 remarks	57 remarks (75%)	4 remarks (7%)	7 remarks (12%)	7 remarks (12%)
	(0.17)	**Lowest:**	1.13		91 remarks	77 remarks (85%)	2 remarks (3%)	11 remarks (14%)	5 remarks (7%)
3	1.93	**Highest:**	2.52	0.95	95 remarks	82 remarks (86%)	2 remarks (2%)	10 remarks (12%)	15 remarks (18%)
	(0.21)	**Lowest:**	1.57		108 remarks	96 remarks (89%)	1 remark (1%)	14 remarks (15%)	10 remarks (10%)
4	1.99	**Highest:**	3.11	2.11	71 remarks	57 remarks (80%)	0 remarks (0%)	3 remarks (5%)	2 remarks (4%)
	(0.52)	**Lowest:**	1.00		86 remarks	74 remarks (86%)	4 remarks (5%)	5 remarks (7%)	16 remarks (22%)
5	2.22	**Highest:**	2.77	0.92	110 remarks	108 remarks (99%)	9 remarks (8%)	12 remarks (11%)	18 remarks (17%)
	(0.19)	**Lowest:**	1.85		101 remarks	98 remarks (97%)	8 remarks (8%)	14 remarks (14%)	13 remarks (13%)
						Mean highest: 82%	**Mean highest:** 5%	**Mean highest:** 10%	**Mean highest:** 13%
						Mean lowest: 89%	**Mean lowest:** 4%	**Mean lowest:** 12%	**Mean lowest:** 14%

HRV = Heart Rate Variability
LF\HF-ratio = Low Frequency/High Frequency ratio; higher score indicates higher stress
Highest = 20 minutes of highest stress measured in teacher
Lowest = 20 minutes of lowest stress measured in teacher
Δ LF/HF-ratio = Highest LF/HF-ratio minus Lowest LF/HF-ratio
NTS = Non-technical skills

RESULTS

Descriptive data

All five surgical teachers were male with a mean age of 51.8 years. Teachers had a mean of 14.6 years of surgical teaching experience. All had completed one or more teach-the-teacher courses. One smoked, none reported use of medication.

The mean age of the ten corresponding trainees, five males and five females, was 36.1 years. Five were surgical residents (one PGY5, four PGY6 residents; the Netherlands has a six-year resident program). The other five were consultant surgeons with a mean experience of 3.8 years. None of them reported to smoke, use medication or oral anti conception. One was pregnant (35 weeks). Teachers and trainees came from different hospitals from all over the country.

Heart Rate Variability data of teachers

Individual overall LF/HF-ratio mean scores ranged from 1.42 to 2.92 (Table 2). The overall mean for the highest LF/HF-ratio value was 2.74 (individual scores ranged from 1.68 to 3.60). The overall mean for the lowest LF/HF-ratio value was 1.61 (individual scores ranged from 1.0 to 2.50). The overall mean difference between the highest and the lowest LF/HF-ratio values was 1.13 (individual differences ranged from 0.55 to 2.11). Individual LF/HF-ratio data per teacher is depicted in Table 2. One teacher's lowest LF/HF-ratio value was 2.50 (Teacher 1), which was equivalent to one teacher's highest value (2.52; Teacher 3), and even higher than one teacher's highest value (1.68; Teacher 2 - Note that only 33% of this data was suitable for analysis).

Coaching behavior of teachers

Nine hundred twenty-two coaching remarks were analyzed: 453 (49%) in the highest LF/HF-ratio condition and 469 (51%) in the lowest LF/HF-ratio condition. For the highest LF/HF-ratio condition 377 (83%) remarks were relevant for coaching and for the lowest LF/HF-ratio condition 419 (89%) remarks. Irrelevant coaching remarks consisted mainly out of listening responses which could not be coded otherwise. Nineteen (5%) of the 377 relevant coaching remarks in the highest LF/HF-ratio condition were directed on non-technical skills (the remainder on main categories A, C-E; Table 1), 39 (10%) were feedback skills (the remainder goal categories I, IV-VII; Table 1) and 53 (14%) were highly detailed coaching (the remainder were content categories 1 and 2; Table 1). Seventeen (4%) of the 419 relevant coaching remarks in the lowest LF/HF-ratio condition were directed on non-technical skills, 50 (12%) were feedback and 58 (14%) were highly detailed.

Descriptive statistics per individual teacher for the highest and lowest LF/HF-ratio condition are shown in Table 2. In the highest LF/HF-ratio condition the mean percentage of total relevant coaching for all five teachers was 82.41% compared to 89.15% in the lowest LF/HF-ratio condition (Table 2 and Table 3). Of the relevant coaching under the highest LF/HF-ratio, a mean of 4.65% was directed on non-technical skills compared to 3.98% in the lowest LF/HF-ratio condition, a mean of 10.09% was feedback compared to 11.60% in the lowest LF/HF-ratio condition and a mean of 13.16% was highly detailed coaching compared to 14.14% in the lowest LF/HF-ratio condition. None of the differences were statistically significant (Table 3). Teachers applied similar coaching behaviors in the highest and in the lowest LF/HF-ratio conditions.

Table 3. Mean coaching for all 5 surgical educators

		Highest LF/HF-ratio	Lowest LF/HF-ratio
Relevant coaching in total coaching	Mean %	82.41%	89.15%
	SD	10.32	4.81
NTS coaching in relevant coaching	Mean %	4.65%	3.98%
	SD	3.40	2.82
Feedback in relevant coaching	Mean %	10.09%	11.60%
	SD	2.91	3.84
Highly detailed coaching in relevant coaching	Mean %	13.16%	14.14%
	SD	5.84	6.16

NTS = Non-technical skills
LF/HF-ratio = Low Frequency/High Frequency ratio
SD = Standard Deviation

LF/HF-ratio graph of the teacher-trainee duo – period 1

For the two 30-minute selected periods of surgery performed by the teacher-trainee duo of team 3, the blindly collected field notes were synchronized with the LF/HF-ratio graph and qualitatively analyzed. In the first period (Figure 1; period 1), the trainee's LF/HF-ratio (red line) decreased to its lowest point. The teacher was explaining the anatomy until the first injury was inflicted (a; stab wound to the femoral vein). The steep increase of the trainee's LF/HF-ratio occurs around the time of the first injury (a) and continues during the second injury (b; 5 bullet wounds to the liver and bowels). No more injuries were inflicted during the remainder of the first period. The second decrease of the trainee's LF/HF-ratio (c) could not be related to any observed events. The teacher communicated calmly, guided the trainees through asking questions, observed and assisted by packing and holding organs aside during the entire period. Although the teacher's LF/HF-ratio also shows increases around the first (a) and second injury inflicted (b), the teacher's graph is much more capricious.

LF/HF-ratio graph of the teacher-trainee duo – period 2

During the second period (Figure 1; period 2) the trainee's LF/HF-ratio (red line) is generally higher than during the first period. The steep increase of the trainee's LF/HF-ratio occurs around the time of the stab wound to the heart (d) that resulted in a mayor bleeding. During the

increase of the trainee's LF/HF-ratio the anesthetist announced that the 'patient' generates no output anymore (e). In the time between d and e the teacher poses two remarks which could be a sign of stress ("Really? You pull the knife out?!" and "Now we really have a problem!"), however, this is not reflected in the teacher's LF/HF-ratio which shows a decrease (blue line). The teacher takes over (e) and shows a slight increase in LF/HF-ratio, however, compared to other moments, it remains fairly low. The decrease of the trainee's LF/HF-ratio coincides with the moment the animal is given up and eventually deceases after an attempt for resuscitation (f). The team continued the operation on a deceased animal. The second steep increase of LF/HF-ratio in both the teacher and the trainee (g) could not be related to any events. The same holds for the second steep decrease of the trainee's LF/HF-ratio (h). The teacher generally remained calm and stood close to the operating table.
Over both periods, there appears to be a relation between the injuries inflicted and an increase of LF/HF-ratio measured in the trainee, but not in the teacher.

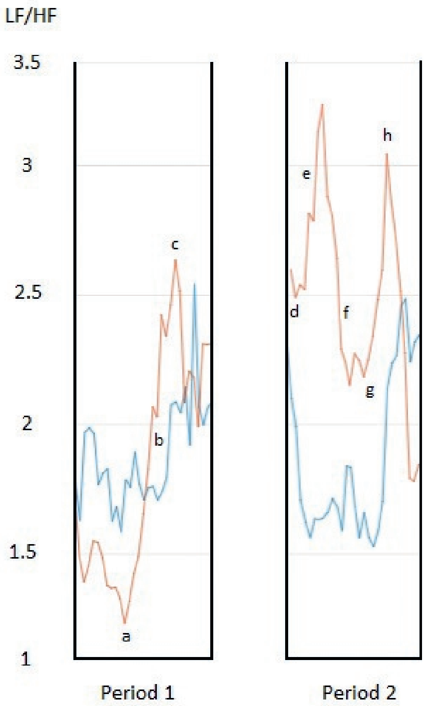


Figure 1. LF/HF-ratio graph of the teacher (blue) and trainee (red; team 3) for time period 1 (14.00 – 14.30) and time period 2 (16.00 – 16.30)

DISCUSSION

We developed an easy to implement, objective, unobtrusive and continuous method to detect different levels of stress in surgical teachers and investigate the coaching behaviors they applied during moments of high and low stress in a highly realistic simulated operation. We also obtained an impression of how stress evolved over time to the minute and the relation to time specific events in the OR. Relating continuous objective stress levels to coaching behavior using a simple wearable sensor is innovative and feasible in a highly realistic acute trauma surgery simulation. Other research to how stress develops over time in surgeons have used saliva samples, which is complex, necessitates interruptions and uses larger time windows.[9]

We had expected that teachers under high stress would have given less non-technical skills coaching, less feedback and less detailed explanations in comparison to low stress. However, regardless of stress, surgical teachers applied coaching behaviors equally frequent. Possible differences, if any were present, may have gone unnoticed due to an overall limited coaching on non-technical skills. The minimal focus of surgical teachers on non-technical skills coaching is in line with previous findings,[13,14] even with educational efforts to increase the frequency of non-technical skills coaching.[23] A predominance of technical skills coaching with directive instructions and minimal feedback is also found during surgical skills training in elective surgical procedures[16,27,28]. This may question the setting of a highly realistic simulation or real operation to study a relation between stress and non-technical skills coaching behavior.

We studied one teacher-trainee duo to relate teacher and trainee stress to salient events and activities in the OR. Although not all increases and decreases in stress could be related to salient events and activities, times of injuries inflicted during the DSATC's integrated acute surgery team training coincided with an increase of stress of the trainee. We cannot rule out that the increases represent 'positive' stress. Some stress can be beneficial and improve focus and efficiency.[6] However, since the increases were large, steep and persisted after the trainee had repaired the injury, we considered the findings indicative for negative stress. Considering the limited time series available for analysis we could not identify stress reducing trainer-trainee interactions as previously observed in time series of stress in elective surgical procedures using the same patch.[20]

We found no corresponding stress in the teacher, likely explained by the fact that the teacher inflicted the injuries himself and was prepared for following events. Another explanation is a well- developed positive stress coping strategy.[7] This teacher (Table 2, number 3) scored average regarding maximum and minimum stress values in a group with marked differences in absolute (increases in) stress values. For example, the lowest stress value of teacher 1 was higher than the highest stress value of teacher 2. Findings may suggest differences in stress coping between individuals as reported by behavioral scientists.[29] When confronted with the same stressors, individuals experience these stressors differently, show different reactions, and apply different coping behaviors. Being able to apply positive stress coping strategies is less likely for the inexperienced trainee than for the experienced teacher,[7] which may explain the on average higher absolute stress of the trainee in period 2 particularly (Figure 1).

This study has several limitations. We used sympathetic and parasympathetic activity (LF) and vagal modulation (HF) ratio as a proxy for stress.[20] The interactions between the sympathetic and parasympathetic nervous system are complex and can also be influenced by factors like physical exercise related respiratory rate and heart rate.[30] However, the subjects in our study remained fairly stationary during operating and physical effort was low. Furthermore, we collected our data in a simulated setting in which the surgical teachers inflicted all injuries which the trainees had to repair. Despite the use of a realistic life animal model, this may have limited the stress experienced by the surgical teachers and the effect on their coaching behaviors. We also collected an important part of our data through video and audio recordings. This may have stimulated teachers to coach to the best of their abilities, despite any stress they experienced. However, then we would have expected more coaching on non-technical skills, since this was an explicitly stated goal of the course. Finally, the data of four out of nine teachers could not be analyzed. The four teachers could have been the ones most vulnerable to stress. We consider this selection bias low because of the mainly technical reasons for drop out. In a prior study we lost 20 percent of participants' data due to loss of skin contact of the patch and connection problems.[20] Improvements in the patch adhesion and sufficient WIFI coverage may reduce this data loss.

Wearable sensors measuring a proxy for stress enable large-scale research of stress in surgical teachers and trainees in an objective, uncomplicated and continuous manner, both during simulated training and training in the operating room. This enables stress to be investigated in future research in relation to teachers' coaching behaviors, trainees' performance, teachers' and trainees' stress coping strategies and personality traits like impulsiveness or stress resistance.

ACKNOWLEDGEMENTS

1. We would like to thank Marja Bulte - ter Meer for her contributions to preparing the research data.
2. The Vital Connect HealthPatch™ (Campbell, CA, USA) we used in this study is further developed into and succeeded by the VitalPatch™ (Campbell, CA, USA). As far as we know, the algorithms forming the basis of our LF/HF-ratio calculations remained the same.

REFERENCES

1. Agha RA, Fowler AJ. The role and validity of surgical simulation. International surgery, 2015. 100(2): p. 350-357.
2. Arora S, et al. Factors compromising safety in surgery: stressful events in the operating room. The American Journal of Surgery, 2010. 199(1): p. 60-65.
3. Arora S, et al. Managing intraoperative stress: what do surgeons want from a crisis training program? The American Journal of Surgery, 2009. 197(4): p. 537-543.
4. Arora S, et al. The impact of stress on surgical performance: a systematic review of the literature. Surgery, 2010. 147(3): p. 318-330. e6.
5. Hull L, et al. The impact of nontechnical skills on technical performance in surgery: a systematic review. Journal of the American College of Surgeons, 2012. 214(2): p. 214-230.
6. Wetzel CM, et al. The effects of stress on surgical performance. The American Journal of Surgery, 2006. 191(1): p. 5-10.
7. Anton NE, et al. What stress coping strategies are surgeons relying upon during surgery? The American Journal of Surgery, 2015. 210(5): p. 846-851.
8. Cochran A, Elder WB. Effects of disruptive surgeon behavior in the operating room. The American Journal of Surgery, 2015. 209(1): p. 65-70.
9. Flinn JT, et al. The effect of stress on learning in surgical skill acquisition. Medical teacher, 2016. 38(9): p. 897-903.
10. McKendy KM, et al. Perioperative feedback in surgical training: A systematic review. The American Journal of Surgery, 2017. 214(1): p. 117-126.
11. Timberlake MD, et al. What do we know about intraoperative teaching?: A systematic review. Annals of surgery, 2017. 266(2): p. 251-259.
12. Mazer LM, et al. Evaluating Surgical Coaching: A Mixed Methods Approach Reveals More Than Surveys Alone. Journal of surgical education, 2018.
13. Alken A, et al. Coaching during a trauma surgery team training: perceptions versus structured observations. The American Journal of Surgery, 2015. 209(1): p. 163-169.
14. Alken A, et al. Feedback activities of instructors during a trauma surgery course. Am J Surg, 2013. 206(4): p. 599-604.
15. Blom EM, et al. Analysis of verbal communication during teaching in the operating room and the potentials for surgical training. Surg Endosc, 2007. 21(9): p. 1560-6.
16. Chen XP, et al. A taxonomy of surgeons' guiding behaviors in the operating room. Am J Surg, 2015. 209(1): p. 15-20.
17. Hauge LS, Wanzek JA, Godellas C. The reliability of an instrument for identifying and quantifying surgeons' teaching in the operating room. Am J Surg, 2001. 181(4): p. 333-7.
18. Roberts NK, et al. Capturing the teachable moment: a grounded theory study of verbal teaching interactions in the operating room. Surgery, 2012. 151(5): p. 643-50.
19. Sutkin G, Littleton EB, Kanter SL. How surgical mentors teach: a classification of in vivo teaching behaviors part 1: verbal teaching guidance. J Surg Educ, 2015. 72(2): p. 243-50.
20. Weenk M, et al. Stress measurement in surgeons and residents using a smart patch. Am J Surg, 2017.

21. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. Bmj, 2000. 320(7237): p. 745-749.
22. Sharma N and Gedeon T. Objective measures, sensors and computational techniques for stress recognition and classification: A survey. Computer methods and programs in biomedicine, 2012. 108(3): p. 1287-1301.
23. Alken A, et al. Integrating technical and non-technical skills coaching in an acute trauma surgery team training: Is it too much? The American Journal of Surgery, 2017.
24. Boffard K. Manual of Definitive Surgical Trauma Care. 1 ed. 2003, London: Edward Arnold.
25. Bleeker CP. Definitive Anaesthetic Trauma Care Courses. [cited 2016 2016-11-7]; Available from: http://www.iatsic.org/DATC.html.
26. Unknown. Definitive Surgical and Anaesthetic Trauma Care Course [cited 2016 2016-11-7]; Available from: http://www.dstc-datc.nl/media/192349/folder_dsatc-datc_2016.pdf.
27. Hu YY, et al. Complementing Operating Room Teaching With Video-Based Coaching. JAMA Surg, 2017. 152(4): p. 318-325.
28. Pernar LI, et al. Exploring the Content of Intraoperative Teaching. J Surg Educ, 2016. 73(1): p. 79-84.
29. Henckens MJAG, et al. Interindividual differences in stress sensitivity: basal and stress-induced cortisol levels differentially predict neural vigilance processing under stress. Social Cognitive and Affective Neuroscience, 2016. 11(4): p. 663-673.
30. Shaffer F, Ginsberg J. An overview of heart rate variability metrics and norms. Frontiers in public health, 2017. 5: p. 258.

6

CHAPTER 7

Evidence-based effective teaching behaviors for complex psychomotor skills training

Alexander Alken[1], Jan-Maarten Luursema[1], Lucas Thornblade[2], Xiaodong (Phoenix) Chen[3], Louise Hull[4], Karen Horvath[2], Harry van Goor[1], Cornelia Fluit[5]

[1]Department of Surgery, Radboud University Medical Center, Nijmegen, The Netherlands
[2]Department of Surgery, University of Washington, Seattle, USA
[3]Department of Surgery, Brigham and Women's Hospital, Boston, USA
[4]Centre for Implementation Science, King's College, London, UK
[5]Radboudumc Health Academy, Radboud University Medical Center, Nijmegen, The Netherlands

Submitted

ABSTRACT

BACKGROUND: Research into teaching in the operating room (OR) is extensive. However, evidence relating surgical teachers' teaching behaviors to the objective acquisition of complex surgical psychomotor skills in trainees is limited.

AIM: The aim of this study is to identify evidence-based teaching behaviors applied in hands-on training which are associated with the increased acquisition of skills in surgical and non-surgical trainees.

METHODS: The databases MEDLINE, PsycINFO and ERIC were searched for relevant papers published before February 14, 2018. Due to comparable training characteristics to complex surgical skill acquisition, papers on sports and music training were also included. Paper screening took place after training sessions to become familiar with the inclusion and exclusion criteria. Inter-rater reliability was determined. Data was extracted and the quality of studies was assessed with the Medical Education Research Study Quality Instrument (MERSQI) and the Newcastle-Ottawa Scale-Education (NOS-E). The impact levels of the studies were also determined.

RESULTS: 18,337 references were identified. Seven studies met criteria for inclusion. Teaching behaviors identified included feedback, instruction, active trainee involvement and demonstrations, and were shown to improve objective skills acquisition in trainees. Feedback and instruction with an external focus on the task and effect were supported by the strongest evidence. There was also significant evidence regarding negative effects of harshly criticizing and belittling teaching behaviors, emphasizing a safe training environment. The data quality and evidence for most teaching behaviors was weak with low levels of impact.

CONCLUSIONS: The identified teaching behaviors of feedback, instruction, active trainee involvement and demonstrations appear important for the hands-on teaching of complex skills but lack strong evidence supporting their effectiveness. Future research should be directed on investigating clearly defined teaching behaviors and their relation to the objective acquisition of skills in surgical trainees.

PRACTICE POINTS

- Although there are many conceptual frameworks for effective teaching in the OR, very little research has been conducted to substantiate the effectiveness of teaching behaviors on the objective development of complex psychomotor skills of trainees in hands-on settings.
- Only one study focused on surgical training. We were surprised by the lack of research since teacher training in effective teaching is considered highly important to establish an effective trainee development.
- Externally focused feedback and instructions making trainees focus on the task and effect instead of how to move had the strongest evidence to support effectiveness in acquiring complex psychomotor skills.
- Behaviors of feedback and instruction, active involvement of trainees and demonstrations have been associated with improved complex psychomotor skill acquisition in trainees. A safe training environment without harsh and belittling criticism is also important.
- The evidence for the effectiveness of these behaviors is mostly weak and limited to improved learner outcomes in a training setting. Future high-quality research with higher levels of impact is necessary.

INTRODUCTION

Skills training is inseparably linked to surgical practice.[1] Although simulation offers a safe training environment[1,2] teaching in the operating room (OR) remains central to surgical education.[1-6] While being taught by experienced surgeons, it is in the OR that trainees develop a set of skills, tacit knowledge, and abilities to cope with complex, stressful and variable situations.[6] Training trainees in the OR is a challenging task because teaching and safe surgical care have to be combined. To offer the best possible training it is important to understand how to successfully support learning in the OR. This may be even more important since the time and opportunities available for teaching in the OR are decreasing.[7-10]

The literature on teaching in the OR is extensive. Behaviors like providing feedback, instructions, explanations, demonstrations, asking questions, and using humor have been identified[11-17] and evaluated.[18-21] Novice trainees need these behaviors to be applied intensively, but as trainees progress, the intensity should be reduced and increasingly substituted by reflective feedback.[22] Although many conceptual frameworks for teaching in the OR are available, most research relies on trainees' perceptions of satisfaction or expert opinions.[23] Research into what, if any, specific teaching behaviors contribute to the actual, objective acquisition of skills in trainees seems scarce. Studies on simulation training are also extensive,[2,24-26] however, it is again hardly addressed which specific hands-on teaching behaviors are used. This makes it difficult to draw firm conclusions and make evidence-based recommendations on how to teach in the OR from a behavior standpoint.

In order to address this evidence gap, we conducted a literature review study according to systematic review principles. Our aim was to identify hands-on, evidence-based teaching behaviors which have been shown to be effective for complex psychomotor skills acquisition in adult trainees. We looked for teaching behaviors that were associated with an objectively measured improvement of psychomotor skills, not a perceived improvement. By gaining more knowledge on this topic and discussing our findings, we expect to be able to express recommendations for surgical teachers to help them teach complex surgical psychomotor skills.

METHODS

Scoping search

A scoping search showed that research on evidence-based teaching behaviors for hands-on surgical complex psychomotor skills was scarce. We considered a skill complex if it was executed with specialized equipment, involves multiple actions and requires conscious cognitive demands. Previous research on surgical skills teaching turned to the fields of sports and music education due to similarities in complexity and training intensity.[8,22] Keeping with this approach, we searched for teaching behaviors in the fields of surgical, medical, sports, and music education.

Search terms

Together with a librarian specialized in systematic reviews we defined our search strategy and searched studies published prior to February 14, 2018, in the MEDLINE, PsycINFO and ERIC databases (Table 1). All references were imported into Endnote X7 (Thompson Reuters, Philadelphia, PA, USA).

Table 1. Search conducted including search terms.

1	((educat* or teach* or instruct* or tutor* or guid* or coach* or train* or feedback or mentor* or supervis* or pedagog* or faculty) adj3 (behavio* or activit* or practice* or interaction* or action* or characteristic* or strateg* or approach*)).mp. [mp=abstract, title, heading word, identifiers]
2	((train* or student* or learn* or apprent* or intern or interns) adj3 (perform* or develop* or progress* or achiev* or competen* or advanc* or enhanc* or improv* or accomplish* or expert* or proficien* or skill* or gain* or grow* or outcome* or effect* or abilit*)).mp. [mp=abstract, title, heading word, identifiers]
3	(surg* or medic* or sport* or athlet* or music* or instrument* or hospital* or operating room* or operating theat* or intraoperative or intra operative).mp. [mp=abstract, title, heading word, identifiers]
4	1 and 2 and 3
5	limit 4 to yr="1995 -Current"

Screening

All references were screened on title and abstract according to a three-stage pre-screening and screening approach (Figure 1). For pre-screening, a medical student was trained using 3 sets of 100 references, independently screened by the student and one author (SA) achieving an inter-rater reliability of Cohen's Kappa 0.7. The student pre-screened all references using inclusion and exclusion criteria (Table 2). We took a conservative approach: in the case of any doubt, a reference was forwarded to the screening stage.

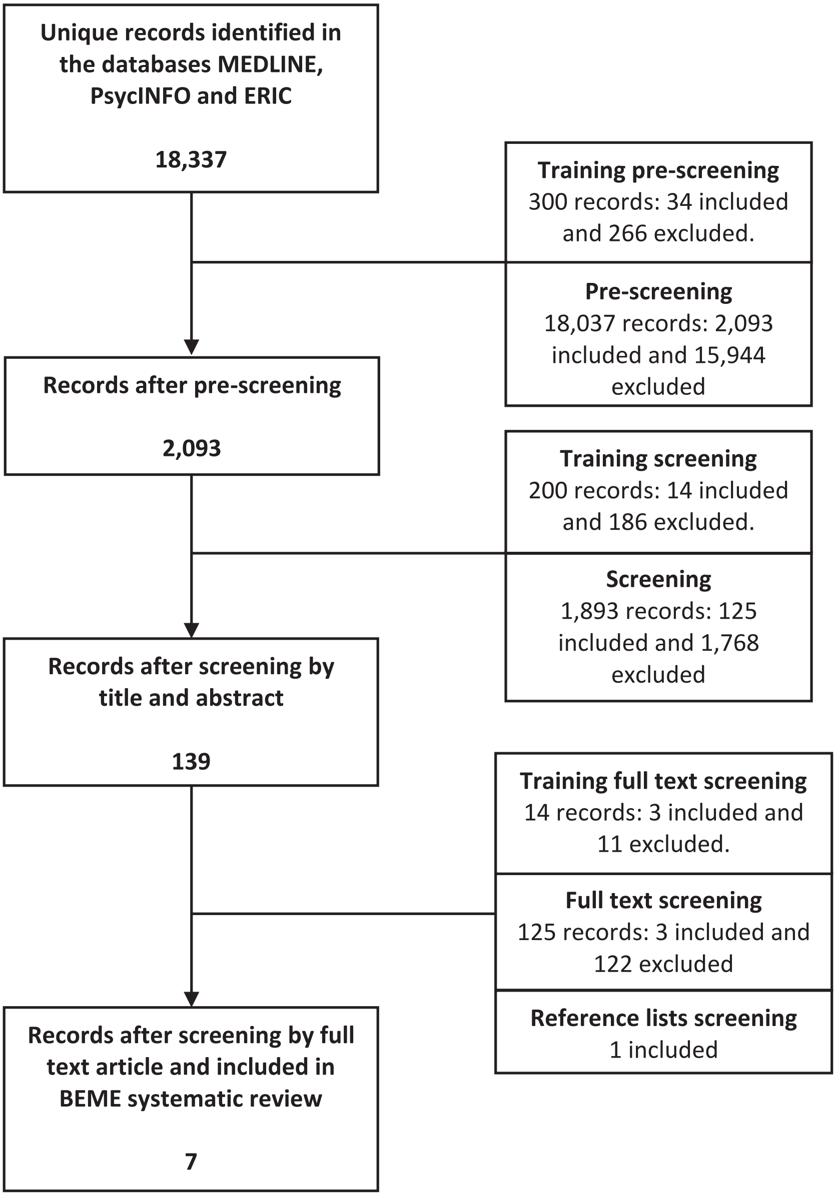


Figure 1. Selection of the reviewed articles.

Table 2. Inclusion and exclusion criteria used during the pre-screening stage.

Inclusion	1	Studies that investigate one or more teaching behaviors in hands-on training situations to assess the effects of these behaviors on the trainees' *complex psychomotor skills** development.
Exclusion	1	Studies in which medical education, surgical education, music education or sports education is NOT the subject of research;
	2	Studies in which psychomotor motor skills are NOT the subject of research;
	3	Studies on the education, revalidation or training of clinical populations;
	4	Studies on non-adult populations;
	5	Studies which do not report original research (e.g. reviews, opinion pieces, letters to the editor, etc.);
	6	Studies published before 1995;
	7	Studies in other languages than English.

* We consider psychomotor skills to be complex if 1. specialized equipment is required for their execution, and/or 2. dynamic decision making is needed to select the proper skill to execute, and/or 3. execution of the skill requires conscious attention even after training (e.g. the skill is physically or cognitively challenging).

In the screening phase we applied stricter inclusion and exclusion criteria (Table 3). Two times 100 references were independently screened by two researchers (SA and JML) for training (Figure 1). An inter-rater reliability of Cohen's Kappa 0.7 was reached. All references were then independently screened by SA and JML, achieving a moderate inter-rater reliability of Cohen's Kappa 0.6. Disagreements were resolved by a third researcher (CF) who was informed about the disagreement, but who was blinded for the individual decisions made by the other researchers. Papers that were included by CF were accepted for full text screening.

For full text screening we applied the strict inclusion and exclusion criteria (Table 3). Due to the moderate inter-rater reliability obtained during screening, two extra training rounds with a total of 14 papers selected by SA were organized to ensure agreement on the interpretation of the inclusion and exclusion criteria. All remaining papers were screened by SA on full text. CF and JML each screened half the amount. We only included papers in which the effectiveness of teaching behaviors was analyzed in relation to the objective acquisition of skills in trainees. Reference lists of included papers were screened for relevant references.

Table 3. Additional criteria for the screening stage on title and abstract and full text screening.

Inclusion	1	See Table 2.
Exclusion	1	Studies that investigate the training of simple psychomotor skills;
	2	Studies which do NOT investigate the effects of teaching behaviors on the objective acquisition of skills in trainees;
	3	1-7 (Table 2).

Data extraction, quality assessment and level of impact

For each experiment in each included study, we extracted general information (e.g. authors, field of research), study set-up (e.g. research aim, design), outcome measures (e.g. data collection methods, study results), and the teaching behaviors that were shown to be effective in the trainees' acquisition of skills.

To assess each experiment's quality and risk for bias we used the Medical Education Research Study Quality Instrument (MERSQI) and the Newcastle-Ottawa Scale-Education (NOS-E) for quantitative research.[27] The MERSQI focuses on study design, number of institutions used for sampling, response rate, subjective or objective data collection, validity evidence of the applied instruments, appropriate data analyses, and impact of the outcome measures. The NOS-E focuses on representativeness of the trainees, selection and comparability of a comparison group, likeliness for study retention, and use of blinded assessors. Because of their different focus, the two scoring systems are considered complementary.[27]

For each experiment, individual MERSQI and NOS-E items were scored and total scores were calculated (maximum MERSQI score: 18; maximum NOS-E score: 6), and compared to the normative score of 12.3 (MERSQI) and 3.58 (NOS-E).[27] If an experiment used two methods to analyze data (e.g. two outcome measures), total MERSQI and NOS-E scores were given for each method and mean scores were calculated. All scores were assigned by one researcher (SA) and discussed with a second researcher (CF). A consensus process was applied in case of disagreement.

MERSQI and NOS-E total scores give an indication of the overall quality. However, quality assessment should also take into account the individual MERSQI and NOS-E items.[27]. Applying the MERSQI and NOS-E enabled us to compare and interpret the quality of studies in relation to the normative scores.

We independently determined the impact levels for each experiment based on recommendations for evidence in medical education from Belfield et al. (2001) who based their work on the research of Kirkpatrick.[28] We identified four main impact levels (Table 4).

Table 4. Level of impact distinguished.

☐	**1b. Reaction teachers/coaches** – Opinions/satisfaction regarding the quality and effectiveness of the coaching behavior (e.g. teachers' responses on questionnaires after training)
☐	**1a. Reaction trainees** – Opinions/satisfaction regarding the quality and effectiveness of the coaching behavior (e.g. trainees' responses on questionnaires after training)
☐	**2. Learner outcomes** – Development of skills in learners (e.g. observation and assessment of newly trained skills acquired by trainees in an experimental setting)
☐	**3. Transfer to real practice** – Development of skills in learners in real practice (e.g. observations and assessment of newly trained skills in the real-life operating room)
☐	**4. Healthcare outcomes** – Changes in health care outcomes that can be attributed to the development of skills in trainees (e.g. the complication rate in operating rooms where trainees with newly acquired skills are involved)

Impact levels are ranked from lowest (1) to highest (4).
Studies with an impact level of 1a and 1b were excluded from our review study (inclusion and exclusion criteria; Table 3).

RESULTS

Of 18,337 references, seven studies describing eight experiments met the inclusion criteria (Table 3; Figure 1). Four were conducted in sports education,[29-31] two in music education;[32,33] one in critical care education[34] and one in surgical education.[35] Data from all experiments were derived from quantitative research methods. Six experiments were entirely focused on adult teachers training adult trainees[29-31,33-35], while two included both adult and non-adult trainees (5th and 6th graders;[32] and 15 to 17 year olds[31]).

7

Study quality

Using the MERSQI and NOS-E checklists, we identified sources of bias per experiment (Table 5). Total MERSQI and mean total MERSQI scores (ranging from 11.5-15) were either near or above the normative score of 12.3.[27] Total NOS-E and mean total NOS-E scores for all experiments were below the normative score of 3.58 (ranging from 1-3.5). Only one experiment used a method which achieved a NOS-E score above the normative score (4).[31] The level of impact for all experiments was focused on learner outcomes (Table 4 and Table 5). This enabled us to assess each experiment's quality and interpret the strength of the results.

Research in surgical and medical skills training

Flinn et al. compared the effects of four feedback behaviors on the objective acquisition of complex laparoscopic skills (Table 5).[35] Trainees who received harshly criticizing feedback performed worse than trainees who received encouraging positive feedback, minimal and neutral feedback, and no feedback. Trainees who received encouraging positive feedback did not perform better than trainees who received minimal and neutral feedback, and no feedback. The researchers concluded that it was not positive feedback which improved learning, but harshly negative and threatening feedback which impaired learning.

McSparron et al. analyzed feedback and instruction behaviors of teachers while they were teaching sub-clavian central venous catheter (S-CVC) insertion to a trainee who was instructed to show challenging learning behavior.[34] Subsequently, the teachers' feedback and instruction behaviors were related to the performance of real novice trainees in a consequent training session (Table 5). Positive feedback (interpreted by the researchers as constructive feedback), suggestions as to how to improve, and step-by-step demonstrations were positively correlated. The regular repetition of learning goals was negatively correlated. The researchers concluded that this may be less effective for technical skills teaching.

Research in sports and musical skills training

Duke and Henninger compared two types of feedback in teaching a musical performance skill: corrections focusing on what trainees had to improve versus corrections focusing on what trainees had done wrong (Table 5).[32] No significant differences were found in trainees' acquisition of skills. The researchers concluded that neutrally provided improvement feedback and fault-focused feedback were equally effective.

Wulf et al. compared the effectiveness of two types of feedback and instructions in teaching sports skills, which either externally focused on the task and effect, or internally focused on how to move (Table 5).[30,31] Regarding accuracy of the trained sports skills, trainees whose teacher provided externally focused feedback and instructions performed better during training and retention, regardless of trainee experience, and especially if teachers had provided feedback after every trial. Trainees could also see themselves how accurate they had performed (trainees could derive perceptual feedback). For movement quality, internally focused feedback and instructions were more effective, but only in novice trainees.

Henninger et al. compared the teaching behaviors of music teachers in teaching a wind instrument (Table 5).[33] Active trainee involvement, by stimulating trainees to ask questions, think aloud, etc., was shown to be important to teach effectively.

Harrison et al. compared the effectiveness of volleyball training in which the trainees determined the sequence and speed of progress versus volleyball training in which the teacher made all decisions (Table 5).[29] No clear conclusions on what was more effective could be drawn.

Table 5. Summary and bias risk assessment of the experiments which were included.

Authors	Field	Aim	Design	Evidence based outcome of learning	Bias risk assessment based on individual MERSQI and NOS-E items (Granted score per item/ Maximum potential score per item)	MERSQI total score	NOS-E total score	Impact level
Duke & Henninger 1998	Music education *Soprano recorder*	To compare the effects of improvement focused versus fault focused corrective feedback on learning a music performance skill on a soprano recorder in elementary school and college students.	Non-randomized 2 group study Setting: Half the trainees received improvement and half received fault focused corrective feedback. One of the researchers provided all feedback and intended to keep a neutral tone. Data collection: The researchers assessed the skills acquisition of trainees based on video-recordings through the amount of trials and time a trainee needed to independently perform in a duet. They also observed and noted the amount of corrections on what to improve, corrections on what went wrong and the amount of positive feedback that was provided.	Trainees in the improvement focused corrective feedback condition did not differ from trainees in the fault focused corrective feedback condition regarding number of trials and time in minutes {**A**} needed to perform independently. Observations of the applied teaching behavior (classification of behaviors in 3 predetermined categories by the researchers {**B**}) showed that improvement focused corrective feedback was higher in the improvement focused condition, fault focused corrective feedback was higher in the in the fault focused condition, but that the amount of positive feedback was equal in both conditions.	**A**: MERSQI: Non-randomized 2 group (2/3); Two institutions (1/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Validity evidence N/A; Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (not randomly selected; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1). **B**: MERSQI: Non-randomized 2 group (2/3); Two institutions (1/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); No validity evidence established (0/3); Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (not randomly selected; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1).	**A**: 14.4* **B**: 12 **M**:13.2	**A**: 3 **B**: 3 **M**: 3	2

M = The mean MERSQI and NOS-E total scores: A+B/2

* = In case instrument validity was not applicable and the maximum potential MERSQI score was 15, the MERSQI score was corrected and multiplied with a factor 1.2 in order to obtain a potential MERSQI total score of 18.

Table 5. *Continued.*

Authors	Field	Aim	Design	Evidence based outcome of learning	Bias risk assessment based on individual MERSQI and NOS-E items (Granted score per item/Maximum potential score per item)	MERSQI total score	NOS-E total score	Impact level
Flinn et al. 2016	Surgical education *Laparoscopic cutting task*	To analyze whether an encouraging feedback style results in less subjective stress and better laparoscopic skills in inexperienced medical students than a harshly criticizing feedback style, a quiet observational feedback style or no feedback.	Four group pre and post test (for the results relevant to our study) Setting: Trainees were equally divided into four groups during simulation training, based on order of sign up: a control group (with no feedback); an observation only group (teacher only observing the trainee using minimal and neutral feedback); an encouraging feedback style group (teacher providing positive feedback and acting encouraging), and a harshly criticizing feedback style (teacher providing harsh criticism, being sarcastic and condescending). Trainees were encouraged to perform as quickly and accurate as possible. Trainees were naive to the study purpose Data collection: The time to task completion was recorded and accuracy scores were collected and transformed into performance scores (based on the SAGES FLS scoring system; no further description). Performance scores were compared for the trainees' first trial, last trial, and the mean performance score of all trials (ranging from 4 to 8) in between.	Trainees in all groups improved significantly during training. Trainees who were provided with harshly criticizing feedback behaviors scored lower than the trainees in the other groups on overall performance scores (time and accuracy scores transformed into a performance score based on the SAGES FLS scoring system; no further description {**A**}). Trainees provided with positive and encouraging feedback scored similar to trainees who received minimal and neutral feedback, and trainees who received no feedback.	**A**: MERSQI: Single group pre test post test (1.5/3); One institution (0.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Validity evidence N/A; Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1).	**A**: 13.2*	**A**: 3	2

Harrison et al. 1995	Sport education *Volleyball*	To analyze whether a practice or command coaching style results in better volleyball sub skills performance in low, medium and highly skilled trainees.	Non-randomized 3 group cohort study Setting: Based on a skills pre-test (unclear who took this) trainees were divided in high, medium and low skilled. Training lasted for 19 days. Per skills group, 50% was provided with the command and 50% with the practice coaching style by a professor in Physical education (no further description). Data collection: Data was collected on the pre-test, day 2 and 3 (midterm test), day 8 and 9 (midterm test) and day 17, 18 and 19 (post-test) according to the American Alliance for Health, Physical Education, Recreation and Dance test and the Stanley spike test (ratio correct to incorrect actions; no further description). Trainee performance was videotaped. Four observers (no further description) rated game play and practice trials.	Regarding practice trials, low skilled trainees in the command style condition performed better on volleyball set skills over the course of 19 days (ratio correct to incorrect performances {**A**}) and low skilled trainees in the practice style condition performed better on the volleyball spike skills (ratio correct to incorrect performances {**B**}).	**A** /**B**: MERSQI: Cohort study (2/3); One institution (0.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Content validity and internal validity established, no relationships to other variables (2/3); Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described; 0/1); Comparable control group (0/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1).	**A**: 13.5 **B**: 13.5 **M**: 13.5	**A**: 2 **B**: 2 **M**: 2	2

M = The mean MERSQI and NOS-E total scores: A+B/2

* = In case instrument validity was not applicable and the maximum potential MERSQI score was 15, the MERSQI score was corrected and multiplied with a factor 1.2 in order to obtain a potential MERSQI total score of 18.

Table 5. *Continued.*

Authors	Field	Aim	Design	Evidence based outcome of learning	Bias risk assessment based on individual MERSQI and NOS-E items (Granted score per item/ Maximum potential score per item)	MERSQI total score	NOS-E total score	Impact level
Henninger et al. 2006	Music education *Wind instrument*	To determine whether experienced teachers and inexperienced pre-service teachers have different effects on trainee progress, trainee performance and apply different teaching behaviors when teaching to play a wind instrument.	Single group post test (for the results relevant to our study) Setting: Trainees selected by the teachers were trained. Training was recorded on video. Data collection: Trainee performance was analyzed by independent experts on a 5-point Likert scale. Trainees were then ranked into the 33% best, intermediate and lowest performing trainees. Teacher and trainee behaviors were analyzed by the researchers with an observational instrument.	Trainees whose performance was classified higher (independent observer assessment: low, intermediate, high, calculated according to a 1 item 5-point rating scale score given by independent observers {**A**}) talked more during training (scoring on earlier developed observational instrument by researchers {**B**}). No significant pair-wise effects; no significant correlation.	**A**: MERSQI: Single group post test (1/3); Three institutions or more (1.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); No validity evidence established (incorrect method for internal validity applied) (0/3); Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (0/1); No comparable group (no separate comparison group; no intervention: 0/1); Non-comparable comparison group (no separate comparison group; no intervention: 0/1); Study retention unlikely source of bias (1/1); Blinded assessment (1/1). **B**: MERSQI: Single group post test (1/3); Three institutions or more (1.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Content validity established, no internal validity (incorrect method applied) and relationships to other variables (1/3); Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (0/1); No comparable group (no separate comparison group; no intervention: 0/1); Non-comparable comparison group (no separate comparison group; no intervention: 0/1); Study retention unlikely source of bias (1/1); No blinded assessment (0/1).	**A**: 11.5 **B**: 12.5 **M**: 12	**A**: 2 **B**: 1 **M**: 1.5	**2**

McSparron et al. 2015	Critical medical care education Placing a sub-clavian central venous catheter (s-CVC)	To develop a procedural objective structured teaching exercise (PrOSTE) and analyze procedural teaching behavior which improves trainee performance.	Single group post test Setting: A standardized trainee instructed to show challenging learning behavior was trained in s-CVC on a simulator. Consequently, in a following training session, groups of three novice trainees were trained in s-CVC on a simulator. Data collection: A standardized trainee rated the teaching skills of each teacher on a questionnaire. One of the researchers assessed the novice trainees' s-CVC skills on a checklist.	Higher ratings of a standardized trainee on 3 items (26-item trainee PrOSTE questionnaire; 1-5 {**A**}) related to higher trainee performance in a consequent training session of novice trainees (assessment by trained observer (researcher); S-CVC 34 item questionnaire; 1-5 {**B**}). The opposite was found for 1 item. No overall correlation between A and B was found.	**A**/**B**: MERSQI: Single group pre test post test (1/3); One institution (0.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Content validity established, no internal validity and relationships to other variables (1/3); Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described: 0/1); No comparable group (no separate comparison group; no intervention: 0/1); Non-comparable comparison group (no separate comparison group; no intervention: 0/1); Study retention unlikely source of bias (1/1); No blinded assessment (blinding not described; 0/1).	**A**: 11.5 **B**: 11.5 **M**: 11.5	**A**: 1 **B**: 1 **M**: 1	**2**
Wulf et al. 1999	Sports education *Golf putt*	To analyze and compare the effects of externally and internally focused instruction on learning the golf putt in trainees with no experience.	Randomized controlled trial Setting: Half the trainees received externally and half received internally focused feedback before each set of training trials. One of the researchers provided all instructions for each trainee during training. During retention no instructions were provided. Data collection: One of the researchers scored the trainees' accuracy scores based on hitting target areas.	Trainees with externally focused instruction outperformed trainees with internally focused instruction on accuracy (hitting target areas with predetermined points, given by one of the researchers; 0-5 {**A**}) during training and retention.	**A**: MERSQI: Randomized controlled trial (3/3); One institution (0.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Validity evidence N/A; Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1).	**A**: 15*	**A**: 3	**2**

M = The mean MERSQI and NOS-E total scores: A+B/2

* = In case instrument validity was not applicable and the maximum potential MERSQI score was 15, the MERSQI score was corrected and multiplied with a factor 1.2 in order to obtain a potential MERSQI total score of 18.

Table 5. *Continued.*

Authors	Field	Aim	Design	Evidence based outcome of learning	Bias risk assessment based on individual MERSQI and NOS-E items (Granted score per item/Maximum potential score per item)	MERSQI total score	NOS-E total score	Impact level
Wulf et al. 2002	Sports education Exp. 1 *Volleyball serve*	To analyze and compare the effects of externally and internally focused feedback on learning the volleyball serve in novice and advanced trainees.	4 group cohort study Setting: Half of the novice trainees received externally and half of the novice trainees received internally focused feedback. The same conditions were applied for the advanced trainees. All feedback was standardized, neutrally formulated and provided during training after every fifth trial (50 in total) by one of the researchers. During the retention test no feedback was provided. Data collection: Based on video recordings two independent observers gave movement quality scores according to guidelines.	Trainees with externally focused feedback after each set of trials during training outperformed trainees with internally focused feedback on accuracy (hitting target areas with predetermined points, given by one of the researchers; 0-4 {**A**}) during training and retention. Effects were independent of trainee experience. Trainees with externally focused feedback outperformed trainees with internally focused feedback on movement quality (independent observer scoring on observational instrument based on guidelines; 0-15 {**B**}) during training, regardless experience. During retention, the advanced externally feedback group outperformed the advanced internally feedback group. The novice internally feedback group outperformed the novice externally feedback group.	**A**: MERSQI: Cohort study (2/3); Two institutions (1/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Validity evidence N/A; Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1). **B**: MERSQI: Cohort study (2/3); Two institutions (1/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Content and internal validity established, no relationships to other variables (2/3); Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); Blinded assessment (1/1).	**A**: 14.4* **B**: 14 **M**: 14,2	**A**: 3 **B**: 4 **M**: 3.5	**2**

Wulf et al. 2002	Sports education Exp. 2 *Soccer pass*	To analyze and compare the effects of externally and internally focused feedback on learning the soccer pass with intermittent or constant feedback.	Randomized controlled trial Setting: Trainees all had experience, received either internally or externally focused feedback, either constant (after every trial) or intermittent (after every third trial; 30 in total). All feedback was standardized, neutrally formulated and provided by one of the researchers. During retention no feedback was provided. Data collection: Accuracy scores were noted by one of the researchers based on target areas.	The externally focused feedback group outperformed the internally focused feedback group on accuracy (hitting target areas with predetermined points, given by one of the researchers ; 0-3 {**A**}) during training and retention, regardless feedback intensity. The intermittent internally focused feedback group outperformed the constant internally focused feedback group on accuracy during training and retention. The constant externally focused feedback group outperformed the intermittent externally focused feedback group.	**A**: MERSQI: Randomized controlled trial (3/3); One institution (0.5/1.5); Sampling response rate ≥ 75% (1.5/1.5); Objective data assessment (3/3); Validity evidence N/A; Data analysis beyond descriptive (2/2); Data analysis appropriate (1/1); Outcomes measuring skills acquisition (1.5/3). NOS-E: No representative group of participants (sampling not described; 0/1); Comparable control group (1/1); Allocation of trainees not concealed (1/2); Study retention unlikely source of bias (1/1); No blinded assessment (0/1).	**A**: 15*	**A**: 3	**2**

M = The mean MERSQI and NOS-E total scores: A+B/2

* = In case instrument validity was not applicable and the maximum potential MERSQI score was 15, the MERSQI score was corrected and multiplied with a factor 1.2 in order to obtain a potential MERSQI total score of 18.

7

Strength of the evidence

The evidence for the described teaching behaviors is mostly weak and limited to improved learner outcomes in a training setting (second lowest level of impact; Table 4). We scored the strength of evidence weak for the results and conclusions drawn by Harrison et al,[29] Duke and Henninger,[32] Henninger et al.,[33] McSparron et al.[34] and Flinn et al.[35] Table 5 shows detailed information regarding the MERSQI and NOS-E bias risk assessment.

We scored the strength of the studies conducted by Wulf et al. moderate, providing the strongest evidence of all studies included in this review.[30,31] The total and mean total MERSQI scores are all above the normative score, and the highest of all experiments included in this review. The total and mean total NOS-E scores are all around, but mostly below the normative score (Table 5). Still, important risks for bias remain.

DISCUSSION

Our goal was to identify evidence-based teaching behaviors which improved complex psychomotor skills acquisition in a hands-on training setting, applicable to the OR. Despite abundant research on skills acquisition, only very few studies analyzed the effectiveness of teaching behaviors in terms of objective improvement of skills in trainees: seven of 18,337 studies met the inclusion criteria and only one study focused on surgical skills teaching (in a simulation setting). Most evidence was derived from studies on teaching sports which, however, contains comparable complex psychomotor skills. We identified evidence-based effective teaching behaviors related to feedback, instruction and active trainee involvement.

Feedback and instructions were effective if provided non-threatening[35] and positive,[34] if they contained suggestions as to how to improve and step-by-step demonstrations,[34] and if they made trainees externally focus on the task and effect.[30-31]

Threatening feedback was found to be harmful for trainees' skills acquisition.[35] The importance of non-threatening feedback is supported by surgical review studies.[23,36] Threatening feedback causes stress in trainees,[35] which is considered harmful since it is induced by a factor outside the learning task itself,[37-39] namely the teacher. Interestingly, Flinn et al. did not find any effects of positive feedback,[35] while McSparron et al. did find an increase in trainees' skills.[34] Comparing both studies is difficult since exact definitions of positive feedback remained unclear. Duke and Henninger reported that negatively formulated feedback (focused on faults) and positively formulated feedback (focused on how to improve) were equally effective if provided with a neutral tone of voice.[32] Suggestions as to how to improve were also shown to be effective.[34] This is supported by review studies on intra-operative teaching.[23,36] Whether feedback is effective depends on how it is provided, but also how it is received and interpreted.[40] This is a complex process influenced by variables like tone of voice, intention, non-verbal behavior and prior relationship attributes established. It would be interesting to investigate how these variables exactly relate to the content of feedback messages and the effects on skills acquisition.

Externally focused feedback and instructions on task and effect were superior to improve accuracy and movement quality in sport skills,[30,31] especially if intensely provided in trainees with experience, and in a setting of perceptual feedback.[31] Only novice trainees improved movement quality more with internally focused feedback and instructions on how to move. In surgical teaching, intensely provided feedback and instructions are considered only effective in novice trainees[23,36] and if trainees cannot derive sufficient perceptual feedback.[22,41] The finding that intense externally focused feedback and instructions are effective in trainees with experience and possibilities for perceptual feedback, may shed new light on the role of the focus, content and intensity of instructions and feedback, and the effectiveness for teaching different surgical (sub) skills in relation to trainee experience and perceptual feedback.

Step-by-step demonstrations have also been shown to improve skills.[34] The importance of deconstructing skills into a series of small steps is also stressed in surgical reviews.[23,36,41] It is considered important to only provide trainees with the key instructions as to what to do. Other information may cause cognitive overload and prevent learning through distraction.[41]

Henninger et al. found that actively involving trainees was effective in skills training.[33] Although active involvement was not clearly defined, they considered it important for teachers to know when to direct trainees, when to allow them to talk, ask questions and verbalize actions and thoughts. Surgical review studies support this finding, although their evidence is primarily based on perceptions and not objective measurements.[23,36] Verbalization by trainees is also considered a key step in surgical skills training because gaining insight in trainees' reasoning processes helps to teach effectively.[41]

Interestingly, the Peyton approach,[41] which is often used in surgical skills training, seems to be compatible with the identified evidence-based teaching behaviors: non-threatening, externally focused feedback on task and outcome, instructions, suggestions how to improve, step-by-step demonstrations and active trainee involvement. The effectiveness of this Peyton approach may be improved by the integration of these behaviors. However, the Peyton approach requires one skill to be taught and performed at least four times in a row,[41] which, in our view, questions its compatibility to teaching in the OR.

Strengths and weaknesses of our study

The strength of our review is the focus on the objective measurement of effects of teaching behaviors on the acquisition of complex psychomotor skills. Our review addresses the growing need for optimal OR teaching behaviors as well as objective assessment of training quality and trainees' skill level to assure safe and effective surgery performed by surgical residents. Our research underlines that much can be gained in the field of surgical educational research. Since only one surgical paper was included one might question the applied inclusion and exclusion criteria we applied. However, the small amount of papers in comparable professions and sports may imply that attention for objective outcomes of different teaching behaviors is generally limited.

The teaching behaviors with the strongest evidence originated from sports research, which questions the translatability of our findings to surgery. Differences exist regarding very fine motor skills training (fingers). However, the intensity and extensity of training necessary to reach proficiency is comparable. Also, proven effective methods in sports, like mental skills training in athletes, have been shown applicable to surgery.[22,42,43] A limitation of our review study is the use of the MERSQI and NOS-E instruments which are specifically designed for medical educational research. The items on these instruments, however, were formulated in a way that made them easily applicable to other research fields. This review only selected papers written in English. To our best estimation we excluded only seven out of 18,337 papers because they were written in non-English languages. This makes us believe all relevant research is represented.

Implications for future research and surgical practice

We were surprised to find a lack of research investigating the effects of teaching behaviors on the objective acquisition of surgical skills. Future research should focus on defining teaching behaviors that are effective for the objective development of complex surgical skills in hands-on training settings (measuring impact beyond perceptions of effectiveness). We advise to study the possibilities for and frequency of externally focused feedback and instructions in surgical skills training and the effectiveness of clearly defined feedback messages in respect of differences regarding formulation, tone of voice, intention and non-verbal behavior. Controlled longitudinal studies are required, both conducted in simulated settings with possibilities of standardizing tasks and teaching behaviors, but also in the clinical real-life OR setting. We advocate to start the incorporation of externally focused teaching behaviors into surgical teaching practice, first in simulation training of simple skills, gradually followed by more complex skills in the real-life OR environment.

ACKNOWLEDGEMENTS

We would like to thank Alice Tillema, librarian at the Radboud University Medical Center for her help in defining and conducting the search applied in this review.

REFERENCES

1. Kneebone R. Simulation in surgical training: educational issues and practical implications. Medical education. 2003;37(3):267-77.
2. Sutherland LM, Middleton PF, Anthony A, Hamdorf J, Cregan P, Scott D, et al. Surgical simulation: a systematic review. Annals of surgery. 2006;243(3):291-300.
3. Li MM, George J. A systematic review of low-cost laparoscopic simulators. Surgical endoscopy. 2017;31(1):38-48.
4. De Win G, Van Bruwaene S, Kulkarni J, Van Calster B, Aggarwal R, Allen C, et al. An evidence-based laparoscopic simulation curriculum shortens the clinical learning curve and reduces surgical adverse events. Advances in medical education and practice. 2016;7:357-70.
5. Konge L, Lonn L. Simulation-based training of surgical skills. Perspectives on medical education. 2016;5(1):3-4.
6. Agha RA, Fowler AJ. The role and validity of surgical simulation. International surgery. 2015;100(2):350-7.
7. Reznick RK, MacRae H. Teaching surgical skills--changes in the wind. The New England journal of medicine. 2006;355(25):2664-9.
8. McCaskie AW, Kenny DT, Deshmukh S. How can surgical training benefit from theories of skilled motor development, musical skill acquisition and performance psychology? The Medical journal of Australia. 2011;194(9):463-5.
9. Snyder RA, Tarpley MJ, Tarpley JL, Davidson M, Brophy C, Dattilo JB. Teaching in the operating room: results of a national survey. Journal of surgical education. 2012;69(5):643-9.
10. Anderson CI, Gupta RN, Larson JR, Abubars OI, Kwiecien AJ, Lake AD, et al. Impact of objectively assessing surgeons' teaching on effective perioperative instructional behaviors. JAMA surgery. 2013;148(10):915-22.
11. Hauge LS, Wanzek JA, Godellas C. The reliability of an instrument for identifying and quantifying surgeons' teaching in the operating room. American journal of surgery. 2001;181(4):333-7.
12. Blom EM, Verdaasdonk EG, Stassen LP, Stassen HG, Wieringa PA, Dankelman J. Analysis of verbal communication during teaching in the operating room and the potentials for surgical training. Surg Endosc. 2007;21(9):1560-6.
13. Roberts NK, Brenner MJ, Williams RG, Kim MJ, Dunnington GL. Capturing the teachable moment: a grounded theory study of verbal teaching interactions in the operating room. Surgery. 2012;151(5):643-50.
14. Sutkin G, Littleton EB, Kanter SL. How surgical mentors teach: a classification of in vivo teaching behaviors part 1: verbal teaching guidance. J Surg Educ. 2015;72(2):243-50.
15. Chen XP, Williams RG, Sanfey HA, Smink DS. A taxonomy of surgeons' guiding behaviors in the operating room. American journal of surgery. 2015;209(1):15-20.
16. Alken A, Tan E, Luursema JM, Fluit C, van Goor H. Coaching during a trauma surgery team training: perceptions versus structured observations. American journal of surgery. 2015;209(1):163-9.
17. Alken A, Tan E, Luursema JM, Fluit C, van Goor H. Feedback activities of instructors

during a trauma surgery course. American journal of surgery. 2013;206(4):599-604.
18. Claridge JA, Calland JF, Chandrasekhara V, Young JS, Sanfey H, Schirmer BD. Comparing resident measurements to attending surgeon self-perceptions of surgical educators. American journal of surgery. 2003;185(4):323-7.
19. Levinson KL, Barlin JN, Altman K, Satin AJ. Disparity between resident and attending physician perceptions of intraoperative supervision and education. Journal of graduate medical education. 2010;2(1):31-6.
20. Rose JS, Waibel BH, Schenarts PJ. Disparity between resident and faculty surgeons' perceptions of preoperative preparation, intraoperative teaching, and postoperative feedback. Journal of surgical education. 2011;68(6):459-64.
21. Jensen AR, Wright AS, Kim S, Horvath KD, Calhoun KE. Educational feedback in the operating room: a gap between resident and faculty perceptions. American journal of surgery. 2012;204(2):248-55.
22. White C, Rodger MWM, Tang T. Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. The Obstetrician & Gynaecologist. 2016;18(1):53-63.
23. McKendy KM, Watanabe Y, Lee L, Bilgic E, Enani G, Feldman LS, et al. Perioperative feedback in surgical training: A systematic review. American journal of surgery. 2017;214(1):117-26.
24. Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. The Cochrane database of systematic reviews. 2009(1):Cd006575.
25. Dawe SR, Pena GN, Windsor JA, Broeders JA, Cregan PC, Hewett PJ, et al. Systematic review of skills transfer after surgical simulation-based training. The British journal of surgery. 2014;101(9):1063-76.
26. Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ. A systematic review of skills transfer after surgical simulation training. Annals of surgery. 2008;248(2):166-79.
27. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education. Academic medicine : journal of the Association of American Medical Colleges. 2015;90(8):1067-76.
28. Belfield C, Thomas H, Bullock A, Eynon R, Wall D. Measuring effectiveness for best evidence medical education: a discussion. Medical teacher. 2001;23(2):164-70.
29. Harrison JM, Fellingham GW, Buck MM, Pellett TL. Effects of practice and command styles on rate of change in volleyball performance and self-efficacy of high-, medium-, and low-skilled learners. Journal of teaching in physical education. 1995;14(3):328-39.
30. Wulf G, Lauterbach B, Toole T. The learning advantages of an external focus of attention in golf. Research quarterly for exercise and sport. 1999;70(2):120-6.
31. Wulf G, McConnel N, Gartner M, Schwarz A. Enhancing the learning of sport skills through external-focus feedback. Journal of motor behavior. 2002;34(2):171-82.
32. Duke RA, Henninger JC. Effects of verbal corrections on student attitude and performance. Journal of Research in Music Education. 1998;46(4):482-95.

33. Henninger JC, Flowers PJ, Councill KH. Pedagogical techniques and student outcomes in applied instrumental lessons taught by experienced and pre-service American music teachers. International Journal of Music Education. 2006;24(1):71-84.
34. McSparron JI, Ricotta DN, Moskowitz A, Volpicelli FM, Roberts DH, Schwartzstein RM, et al. The PrOSTE: identifying key components of effective procedural teaching. Annals of the American Thoracic Society. 2015;12(2):230-4.
35. Flinn JT, Miller A, Pyatka N, Brewer J, Schneider T, Cao CG. The effect of stress on learning in surgical skill acquisition. Medical teacher. 2016;38(9):897-903.
36. Timberlake MD, Mayo HG, Scott L, Weis J, Gardner AK. What do we know about intraoperative teaching?: A systematic review. Annals of surgery. 2017;266(2):251-9.
37. Joëls M, Pu Z, Wiegert O, Oitzl MS, Krugers HJ. Learning under stress: how does it work? Trends in cognitive sciences. 2006;10(4):152-8.
38. Schwabe L, Joëls M, Roozendaal B, Wolf OT, Oitzl MS. Stress effects on memory: an update and integration. Neuroscience & Biobehavioral Reviews. 2012;36(7):1740-9.
39. Vogel S, Schwabe L. Learning and memory under stress: implications for the classroom. npj Science of Learning. 2016;1:16011.
40. van de Ridder J, McGaghie WC, Stokking KM, ten Cate O. Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. Medical education. 2015;49(7):658-73.
41. Nicholls D, Sweet L, Muller A, Hyett J. Teaching psychomotor skills in the twenty-first century: revisiting and reviewing instructional approaches through the lens of contemporary literature. Medical teacher. 2016;38(10):1056-63.
42. Anton NE, Stefanidis D. Should surgeons have mental skills training? : Oxford University Press; 2016.
43. Rogers RG. Mental Practice and Acquisition of Motor Skills: Examples from Sports Training and Surgical Education. Obstetrics and Gynecology Clinics of North America. 2006;33(2):297-304.

7

CHAPTER 8

Integrating technical and non-technical skills in hands-on surgical training

Alexander Alken[1], Cornelia Fluit[2], Jan-Maarten Luursema[1], Harry van Goor[1]

[1]Department of Surgery
[2]Radboudumc Health Academy

Radboud University Medical Center, Nijmegen, The Netherlands

Medical and Surgical Education: Past, Present and Future 2018; InTech

ABSTRACT

Safe and effective surgery requires high-quality technical and non-technical skills. Although the importance of non-technical skills has become increasingly clear, today's surgical curricula still lack formal training in non-technical skills. In this chapter, we discuss how to integrate technical and non-technical skills training into surgical curricula and provide strategies on how to teach both skill sets concurrently in a hands-on setting.

INTRODUCTION

The operating room is a hectic and dynamic teamwork environment requiring safe surgical practice. To achieve this, surgeons have to develop highly effective technical and non-technical skills that are both built upon formal training.[1] Since technical and non-technical skills are strongly associated to another[1] and have to be applied concurrently in the real-life operating room, these skills also have to be trained concurrently.[2-5] However, to date, the training of non-technical skills is still not effectively and fully implemented in surgical curricula.[1]

Surgical curricula have always had a strong focus on the development of surgical technical skills, aiming for high-level psychomotor skills, swift eye-hand coordination, and dexterity.[6] Since it became clear that adverse events in surgery are not so much caused by deficiencies and errors in technical skills but rather by deficiencies and errors in non-technical team and communication skills, the attention for non-technical skills training has been increasing.[1,7]

In this chapter, we discuss how to effectively integrate technical and non-technical skills in surgical curricula, focusing on teaching behaviors, training strategies, and simulated and real-life operating room training. We address the current knowledge on acquiring technical skills in the section "How do we learn and teach surgical technical skills?" and non-technical skills in "How do we learn and teach non-technical skills in surgery?". We discuss training settings used for concurrent technical and non-technical skills training in "Environments used for teaching non-technical skills next to technical skills in surgery." We address frameworks that can be helpful to integrate technical and non-technical skills teaching in the section "Effective frameworks for integrated technical and non-technical skills teaching." Finally, we share our view on how to successfully establish a surgical curriculum with integrated technical and non-technical skills training and teaching in "Integrating technical and non-technical skills training during hands-on surgical teaching in a curriculum".

We realize that non-technical skills of surgeons encompass more than communication and team skills, such as clinical decision making and stress coping.[5] In this chapter, the term 'non-technical skills' refers to the social team and communication skills necessary for successful surgery, encompassing effective communication (e.g., sharing information efficiently), collaboration (e.g., assisting other team members when needed), coordination (e.g., timely asking team members for requests), leadership (e.g., directing team members), and situational awareness (e.g., asking for information updates regarding the patient's condition).[8,9]

HOW DO WE LEARN AND TEACH SURGICAL TECHNICAL SKILLS?

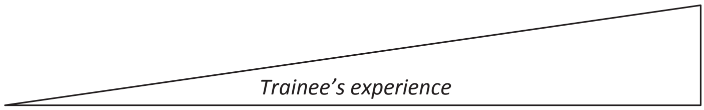
Technical skills are core skills of a surgeon. They refer to all goal-directed psychomotor actions. Handling a scalpel to gain access to a patient's abdomen and handling a needle to repair a ruptured vein are examples of technical surgical skills. We discuss the current knowledge available regarding the learning, training and teaching of technical skills, and surgical technical skills in particular.

To understand how technical skills are acquired, two main principles are important. First, the acquisition of technical skills is not a linear process.[10] At the start of training, trainees are completely unfamiliar and inexperienced. There are many untrained aspects of the skill that can be improved relatively easily resulting in a rapid growth of trainees during their first trials. After the first trials, however, the basics of the skill are acquired and the speed of progress starts to slow down. All aspects of the skill are familiar but now have to be optimized, which requires much more time. The second principle regarding the acquisition of technical skills involves the way a skill is cognitively approached by trainees. In the beginning stages of learning, trainees depend on explicit rules to perform the skill that demands extensive cognitive effort. In later stages, the skill is partially or completely automated and trainees are less dependent on rules. Performing the skill then requires far less cognitive effort and enables trainees to focus on other actions, activities, or aspects in the training or working environment.[10] Both principles are important for designing the training and the teaching that is to be provided.

Distributed, short training sessions with sufficient resting periods in between are most effective to acquire technical skills.[10] Effective training stimulates trainees to engage in deliberate practice, meaning that trainees repetitively train skills in a dedicated and conscious manner according to clear and achievable learning goals. Training of skills has to be simple at the start and complexity should gradually increase over time in accordance to each trainee's individual progress and needs. Teaching has to contain demonstrations, instructions, and immediate feedback intensely in case of novice trainees[10] to prevent the wrong acquisition and automation of skills.[11] But as trainees progress, this intensity should be reduced accordingly and feedback should be increasingly given reflectively (Figure 1).[10] Highly frequent feedback stimulates the initial acquirement of skills in the short term, whereas less frequent feedback seems to stimulate the retention of the learned skills in the long term. If tasks enable trainees to derive perceptual feedback and trainees are able to evaluate and improve their skills themselves, no feedback may be required. Under such circumstances intense feedback can even hamper learning.[10] Feedback has to be provided in a constructive manner.[12] Based on first-hand observations, teachers should provide specific information to trainees regarding what went well and what needs to be improved.[12] All teaching has to occur in a calm, supportive, and respectful way to create a safe and effective teaching climate.[13,14] The use of neutral, non-disturbing language is essential in achieving this.[15] When teaching technical skills, the information density should be limited to how to perform the skill without focusing on other aspects or additive information (e.g., other options and the environment).[11] Such information can prevent trainees from learning because executing the task and simultaneously processing the teaching require trainees to process too much information, which can easily cause cognitive overload. It is also important that trainees can learn by exploring and committing errors to experience firsthand what works best and what possible consequences are.[10] Simulation training plays an important role here.

Effective teaching behavior

Research to effective surgical skills teaching is primarily based on trainees', teachers', and educational experts' perceptions of effective teaching. Furthermore, research relating teachers' teaching behaviors to the actual acquisition of trainees' skills is scarce, especially in surgical and other medical training. This makes it hard to determine evidence-based how teachers exactly should teach. A recently conducted systematic review study on technical skills teaching in medicine, sports, and music found that feedback, instructions, suggestions for improvement, and demonstrations by teachers improved the skills development in trainees.[16]



Timing	Concurrent	Reflective
Frequency	High	Low
Information density	Low	High

Figure 1. Changes in feedback on technical skills according to the trainee's progress.

It was also found important that teachers stimulated trainees to verbalize their thoughts and reasoning processes. A safe training environment was also found to be important. However, it remained unclear how these behaviors should exactly look like in order to effectively improve the acquisition of skills in trainees. Only one behavior was found to be sufficiently supported by evidence and was described elaborately: instructions and feedback that made trainees externally focus on the task and the effect was more effective than instructions and feedback that made trainees internally focus on how to exactly move their body parts. It may be more effective to explain by "When suturing, move the needle in a circular motion as if you are going to make a full circle." than by "When suturing, move your wrist with a turning motion." However, results were only shown in psychomotor sport skills teaching. It is unclear whether it is also effective in surgical hands-on training.

HOW DO WE LEARN AND TEACH NON-TECHNICAL SKILLS IN SURGERY?

Most non-technical skills trainings are built upon (the combination of) three approaches: theory-based, demonstration-based, and simulation-based training.[1] In the theory-based training, trainees are classroom-like taught what non-technical skills are, why they are important, and how they can be applied in the operating room. Demonstration-based training adds demonstrations of non-technical skills to the theory-based approach. Trainees observe, for example, video-recordings of a simulated operation and discuss the behaviors they have seen, possible consequences, and solutions. Both approaches are low cost and easily organized. Both approaches improve trainees' knowledge, awareness, and attitudes. However, there are no possibilities for trainees to apply and train non-technical skills by doing. Simulation-based approaches do have a training-by-doing component. Trainees apply and train non-technical skills hands-on in a safe environment, varying from simple bench-top models to full-scale simulated operating rooms with, for example, a patient simulator. This enables the hands-on training of basic and advanced non-technical skills in a realistic and multidisciplinary team setting. Simulation training is easily compatible with the theory-based and demonstration based approaches. Simulation has been shown to improve skills acquisition, both in training and in the real-life operating room. Drawbacks are that simulation training is costly, demands extensive organization since realistic operating teams have to be composed, and requires guidance by trained teachers. Blended approaches in which theory, demonstrations, and simulations are combined are advocated. Theory and demonstrations should be the focus in the early stages of non-technical skills training. In following stages, non-technical skills should be applied and trained in simulated settings.[1,4] Simulation has been shown to be most effective in acquiring non-technical skills.[2] Non-technical skills training should be long-term, structured according to learning goals, trainees' individual needs and experiences, and firmly embedded into surgical training curricula.[1,4] Multiple and distributed training sessions are considered most effective.[2]

Although simulation training has been shown to be effective within different surgical specialties,[2] the applied interventions are often minimally described.[17] As a consequence, it remains unclear how non-technical skills teaching should exactly look like to be effective.[1] Most commonly used are debriefing sessions immediately after simulation in which reflective feedback on non-technical skills is provided by a teacher.[2] Debriefing sessions are considered highly important for teaching non-technical skills. Trainees receive feedback on their non-technical skills performance, but also reflect on their non-technical skills.[15] Teacher feedback on non-technical skills has to be based on first-hand observations of trainees' performance and has to address specific strong points, weaknesses, and suggestions for improvement. It should always be provided in a respectful way and with a neutral tone of voice. Teachers have to structure their feedback based on observations and evaluations according to non-technical skills assessment tools.[1,15] Training is essential to enable surgical teachers to effectively analyze and teach non-technical skills.[1,2,,15] Extensive training and coaching by non-technical skills experts are required.

ENVIRONMENTS USED FOR TEACHING NON-TECHNICAL SKILLS NEXT TO TECHNICAL SKILLS IN SURGERY

The operating room

To date, surgical trainees generally learn non-technical skills informally and unstructured *as they pass by* in the real-life operating room.[1] Teaching in the operating room is a challenging task in itself since the patient's safety is the most important aspect that often pushes teaching and training to the background.[18] Teaching as it occurs in the operating room is almost entirely focused on technical skills.[15] The same goes for debriefing sessions after the operation.[19] Non-technical skills teaching remains to be under taught[15] and surgical teachers are not sufficiently trained to teach non-technical skills.[1] Non-technical skills are not yet a part of the surgical educational culture and teachers are inclined to avoid it, often unaware. This means that the real-life operating room is not the best place to teach and train non-technical skills, especially when trainees lack technical experience.

Simulation

Simulation offers surgical teachers the possibility to put the teaching of trainees in the forefront.[18] Trainees can train and apply skills and build experience in a safe environment without any risks for patients. For the training of technical skills, synthetic bench top models, box trainers, virtual reality simulators, animal cadavers, human cadavers, and live animal models are commonly used.[20,21] Recently, non-technical skills have slowly been added to the skill sets trained by simulation. Possibilities to integrate non-technical skills next to technical skills training have been developed, for example by using simulated operating rooms with live animal models,[22] manikins, and/or synthetic or hybrid models.[18]

8

Animal model simulation

Using an animal model adds to the reality of the experience in a simulated operating room. Trainees have to deal with real bleedings, time pressure, and hectic teamwork within realistic operating teams.[22] This reality is considered to contribute to the transfer of the acquired non-technical skills to the real-life operating room.[2] However, our research showed that integrating technical and non-technical skills during a training using live animal models remained difficult and resulted in a main focus on the teaching of technical skills and hardly on non-technical skills.[22,23] It was assumed that the use of a live animal made training activities too unstable to properly teach non-technical skills next to technical skills. The animal's condition can really deteriorate without being able to pause the situation. Like in the real-life operating room, the main focus may remain on keeping alive the animal patient and on the technical skills necessary to do so. The teaching of non-technical skills may easily shift to the background. Furthermore, the occurrence of stress caused by real bleedings, for example, is known to cause cognitive overload. This may have further limited the teachers in teaching non-technical skills.[22]

Simulated human patient operation

A simulated operating room equipped with an operating table, equipment, instruments, and synthetic or virtual reality models offers more opportunities for integrating non-technical next to technical skills teaching than other environments.[20] The use of synthetic or virtual models creates the possibility to teach and train surgical procedures and the necessary technical skills in an authentic setting. By adding a realistic operating team, the teaching and training of non-technical skills are possible. Such simulation setting not only offers extensive debriefing possibilities after simulation, for example with the help of video-recordings but also offers the possibility to freeze the condition of the patient in time. Pause and reflect procedures focused on non-technical skills can then follow immediately, even in acute situations.[22] While still in the simulation, trainees receive feedback and have the opportunity to immediately apply and train the improved non-technical skills during the remainder of the simulation. However, no research has yet been conducted to the effectiveness of such pause and reflect procedures.

EFFECTIVE FRAMEWORKS FOR INTEGRATED TECHNICAL AND NON-TECHNICAL SKILLS TEACHING

Many frameworks are available for teaching surgical skills and skills in general. We discuss four approaches that in our opinion are superior at the moment in contributing to the effective integration of technical and non-technical skills teaching in surgery.

The Peyton four step approach

The Peyton four step approach is a widely accepted method to teach and train technical medical and surgical skills (Table 1).[11] During the first step, the teacher shows the trainee how to perform the skill without any instruction. This is to enable the trainee to entirely focus on the performed motor skills without potentially distracting verbal information. During the second step, the teacher shows and explains the skill according to manageable, logically sequenced part-tasks (steps within the skill). Explanations should be limited to only the key information to prevent cognitive and information overload. In the third step, the teacher performs the part-tasks according to the instructions provided by the trainee. This enables the teacher insight into the trainee's understandings and misunderstandings and helps the teacher to adjust the teaching. During the fourth step, the trainee first verbalizes what to do in each part-task (the teacher checks whether the part-tasks are understood) and then executes the skill (the teacher checks whether the skill is performed correctly). Misunderstandings or mistakes have to be immediately corrected to teach the skill correctly and prevent the wrong automation of the skill. Although widely used, the evidence for the effectiveness of the Peyton four step and similar approaches is limited.[11]

Table 1. Peyton four step approach for technical skills.

	Teacher	Trainee
1	Performs	Observes
2	Shows and explains	Observes
3	Performs	Instructs
4	Observes	Explains and shows

The Zwisch model

The Zwisch model guides teachers in tailoring their teaching and granting autonomy in accordance to each individual trainee's level for each surgical procedure.[24] The model's first stage, 'show and tell,' applies to inexperienced trainees. The teacher performs the procedure, shows, and explains how it is done, while the trainee observes and assists. This step looks similar to the second step of the Peyton approach; however, the Zwisch model does also allow teaching beyond the key information (e.g., background information, other options, and information on team skills) if the trainee's level of experience allows this. When the trainee is familiar with the entire procedure and is able to actively assist (e.g., anticipating on the progress of the procedure), the trainee moves on to the second stage, 'smart help.' In this stage, the teacher switches between self-performing and assisting the trainee in performing the procedure. The trainee performs most of the procedure based on the teacher's continuous instructions and feedback. This stage continues until the teacher deems the trainee able to perform the entire procedure with this 'smart help.' During the third stage, 'dumb help,' the trainee performs the procedure entirely while the teacher assists and only provides feedback and instructions for fine-tuning the trainee's skills. The final step is called "no help" during which the teacher monitors the trainee performing the procedure and only provides minimal advice. The Zwisch model is supported by psychomotor learning theories,[24] however, research on the effectiveness is mostly lacking.

The BID model

The three-phase briefing-intraoperative teaching-debriefing (BID) framework[25] has been shown to be effective in structuring the teaching before, during, and after hands-on surgical training in the operating room.[13,14] The briefing phase is characterized by discussing and setting learning goals for trainees to work on during the intraoperative phase.[25,26] Trainees come up with learning goals or teachers do suggestions, for example, based on their prior experiences with the trainee. Generally, trainees are inclined to focus on technical skills[14] so teachers may have to stimulate trainees to formulate non-technical learning goals. Research suggests that if clear learning goals on non-technical skills are lacking the teachers' attention for non-technical skills will be minimal.[22] The teacher and trainee should agree to focus on one or two learning goals,[25] which help the teacher, but also the trainee, to focus on what has to be taught and trained during the intraoperative phase. The briefing generally takes only a few minutes.

The teaching during the intraoperative phase should be aimed at achieving the learning goals and form the guidelines for the teacher to structure the training.[25] How to effectively teach during the intraoperative or the hands-on training phase is discussed in the sections concerning the learning of technical and non-technical skills.

The final phase is the debriefing during which the teacher and trainee reflect on the achievement of the learning goals.[25] Research showed that teachers during the debriefing are inclined to emphasize what went wrong with a focus on technical skills and without sufficient elaboration on how to improve.[14] A teacher-trainee dialog should be established containing honest and specific strong points, weaknesses, suggestions, and solutions for future practice based on the formulated learning goals. Also the underlying schemes why trainees acted the way they acted during the intraoperative phase should be discussed.[25] This provides both the trainee and the teacher insight into the trainee's thoughts, beliefs, and reasoning processes. This offers reflection and deeper learning for trainees and provides teachers with the opportunity to specifically develop or improve trainees' schemes regarding technical, and in particular non-technical skills. Along this process, teachers should preferably ask open questions. Debriefing sessions should always result in individual learning goals for trainees' future performance.[15] The use of video-recordings of the trainee's performance has been shown to be effective in debriefing.[2] It helps and teaches trainees to reflect. Mistakes can be discussed and remediated.

The 4 C/ID model

The 4 component instructional design (4 C/ID) model is specifically designed for the teaching of complex skills (Figure 2)[27,28]. The model distinguishes four components for designing a successful training program: learning tasks, supportive information, procedural information, and additional part-task training. The learning tasks are at the core of the 4 C/ ID model and encompass the training of whole-task procedures in an authentic and realistic training setting *from the very start*. Whole-task procedures require trainees to combine knowledge, skills, and attitudes and enable them to train these aspects in a realistic relation to each other. This is an important contrast to classical learning theories, which generally prescribe the division of complex tasks into subtasks for training. The 4 C/ID model considers such an approach only effective for the learning of simple skills or skills that can be automated through training. Obviously, confronting a trainee with a new, complex task at the very start of training will induce cognitive overload and hinder learning. To prevent this, the whole-task training should start with the simplest or most simplified version of this task. This requires much less information processing, reasoning, and problem-solving. As trainees improve, they develop schemes and automate skills that then require less cognitive resources and enable them to focus on other aspects of the skill. As trainees' capabilities improve, the complexity of training should gradually increase accordingly. Whole-task training will improve the effective transfer of the trained skills to real practice, which also requires the whole task to be performed.[27,28] Regarding the training of technical and non-technical skills, the 4 C/ID model would require trainees to train these skills concurrently from the very start. Training should gradually progress from the simplest to the most complex version of a surgical procedure.

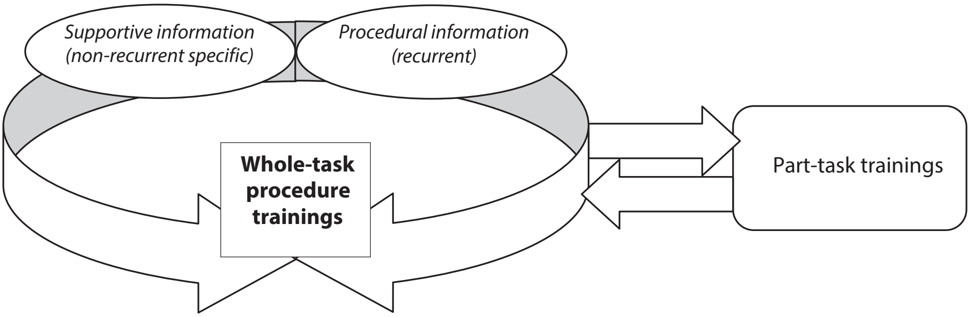


Figure 2. The 4 components of the 4 C/ID-model.

Supportive information should be available for trainees during training.[27,28] Supportive information encompasses knowledge that trainees need for reasoning and problem-solving during non-recurrent, situation-specific aspects of the learning task, for example, information on consequences based on patient-specific characteristics (technical skills) or operating team composition (non-technical skills). Procedural information encompasses knowledge of recurrent procedural aspects of the learning task.[27,28] Preferably, this information is provided to trainees during training when the situation requires it, for example, conducting a physical test on the patient (technical skills) or working with a checklist (non-technical skills) during the operation. Supportive information and procedural information encompass the teaching activities applied by the teacher. The final component of the 4 C/ID model is additional part-task training.[27,28] Although whole-task training is key and remains the focus during the entire training program, some skills require a high level of automation and additional training. Such skills can be trained and automated in separate training sessions.

The development of and research to integrated technical and non-technical skills training in surgical specialties are very scarce.[5] Currently, modules are focused on developing technical skills or non-technical skills, but barely on integrating and developing both skill sets concurrently. Although the 4 C/ID model is built upon solid learning principles and theories, to our knowledge, its effectiveness has not yet been investigated in relation to surgical teaching.

Nevertheless, we believe this model helps to integrate technical and non-technical skills teaching successfully.

INTEGRATING TECHNICAL AND NON-TECHNICAL SKILLS TRAINING DURING HANDS-ON SURGICAL TEACHING IN A CURRICULUM

In this section, we put the aforementioned theory together and share our view on how to effectively integrate technical and non-technical skills training and teaching in surgical curricula, using both simulated settings and the real-life operating room. We advocate a surgical curriculum to be organized in different training modules. Each training module is composed out of different training sessions. In these training sessions, the actual teaching and training takes place. We provide recommendations on three main aspects: learning goals and assessment; training modules and training sessions; and teaching and training within training sessions.

Learning goals and assessment

To ensure the formal and structural training of non-technical skills next to technical skills, long term learning goals on non-technical skills have to be formulated next to long-term learning goals on technical skills. Long-term learning goals are formulated by program directors in cooperation with surgical teachers. All have to be achieved by trainees during the curriculum. The long-term learning goals on non-technical skills can be based on research (for example, the studies conducted by Hull et al.[8] and/or Yule et al.)[9] combined with specific needs of the workplace.

Based on the long-term learning goals, the program directors and surgical teachers formulate short-term learning goals (Figure 3). Short-term learning goals have to be achieved by trainees during the training modules within the curriculum. Short-term learning goals address both general and procedure-specific technical and non-technical skills. General learning goals apply to all surgical procedures. Procedure-specific learning goals specifically apply to distinctive surgical procedures.

The formulated short-term learning goals are the guidelines for the surgical teachers and trainees to formulate personal learning goals for each individual trainee. These personal learning goals enable teachers to tailor the teaching to each individual trainee's needs, experience, and interests. Personal learning goals are specific and achievable. They can be achieved in one or in multiple training sessions.

Each trainee's individual performance and progress is assessed and monitored with assessment instruments on technical skills (like the Objective Structured Assessment of Technical Skills (OSATS)) and non-technical skills (like the Observational Teamwork Assessment for Surgery (OTAS)[8] or Non-technical Skills for Surgeons (NOTSS)).[9] The trainees' progress on non-technical skills is structurally analyzed after each training session and documented with the purpose to provide trainees with feedback and personal learning goals and to improve future performance in upcoming training sessions. Each training module finishes with a summative pass or fail test for trainees on technical and non-technical skills concurrently, according to the short-term learning goals.

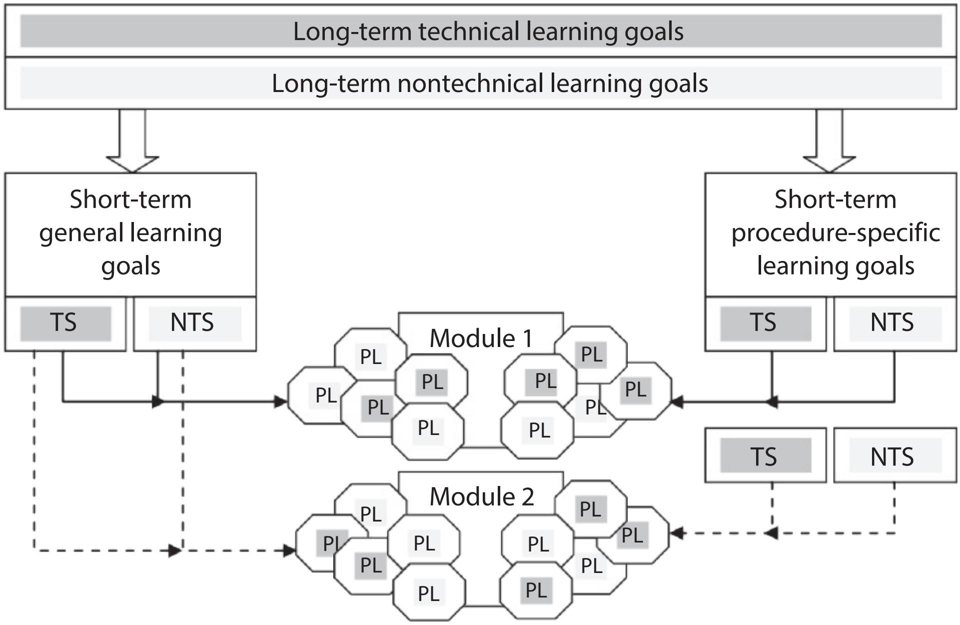


Figure 3. Modular structure of a curriculum for integrated skills teaching and learning.

TS / dark = Technical skills
NTS / light = Non-technical skills
PL = Personal learning goal

Surgical teachers have to be trained and coached in observing, assessing, and teaching non-technical skills by non-technical skills teaching experts. Teachers should regularly reflect on their non-technical and technical skills teaching abilities based on peer observations and feedback. Assessing the teachers' abilities is not a goal in itself but rather a method to strive for perfection, learn from each other, and stay up-to-date regarding effective surgical teaching research and frameworks.

Preferably, non-technical skills are integrated over the entire line of surgical education, starting in undergraduate medical education all the way through to continuing postgraduate surgical education. It may well be that technical skills require more training and teaching effort to develop than non-technical skills. The perfect ratio is not known. However, effective non-technical skills can only be achieved through a formal, structured, and sufficient installation of training possibilities.

Training modules and sessions

We advocate that surgical training curricula are composed out of training modules focused on specific surgical procedures (e.g., a laparoscopic cholecystectomy module, an open inguinal hernia repair module, etc.). What training modules are exactly incorporated in a curriculum is decided by the program directors depending on the relevance of the modules and the intentions of the curriculum. Technical and non-technical skills are integrated from the start of a module so trainees learn that both skills are connected and have to be applied and trained concurrently.

Integrated training sessions

A training module consists out of several distributed training sessions. Two types of training sessions are distinguished: integrated training sessions and focused training sessions. Integrated training sessions focus on the teaching and training of both technical and non-technical skills concurrently within the same training session, in a simulated setting, or in the real-life operating room. If trainees have no prior experience, a training module may typically start with an integrated training session in which trainees purposefully observe and analyze their teachers' technical and non-technical skills (e.g., based on observational assignments and observational instruments) while their teachers are performing simple versions of the procedure in the real-life operating room. The following integrated training sessions are gradually organized, in accordance to each trainee's individual level, progress, and needs, from:

• *Simple or simplified versions to complex versions of the procedure*

Gradually moving from simple to complex ensures that trainees do not experience cognitive overload and can start performing procedures or components of procedures themselves in an early stage.

• *Highly controlled simulated to barely controlled real-life training environments*

Gradually working from a highly controlled to a barely controlled environment ensures there is sufficient room for teaching technical and non-technical skills, especially in the early stages of training. A highly controlled environment is characterized by simulation, which enables pause and reflect procedures and learning from mistakes. A typical barely controlled environment is the real-life operating room. As trainees start to perform skills themselves, they start in a simple simulation setting (e.g., bench top models). Depending on the trainee's progress, simulated environments become gradually less controllable, increasingly realistic and more and more replaced by training in the real-life operating until the point of independent practice is reached.

• *Teaching of high intensity to teaching of low intensity*
The teaching intensity gradually fades from strict guidance and intensive teaching in trainees who are inexperienced to distant observation by the teacher with reflective feedback only when trainees are experienced. The Zwisch model can help teachers to determine the necessary teaching intensity.

• *Low trainee contribution to high trainee contribution*
Trainee contribution gradually increases on the way to independent practice. Trainees first train and apply only the simple components of the procedure, scattered throughout the integrated training sessions. The more complex components are still performed by the teacher. Since full procedure training is important, trainees then purposefully observe and analyze the technical and non-technical skills still performed by their teacher or assist their teacher during the complex components. As trainees progress, next to the already trained components, increasingly more complex technical and non-technical components of the procedure are added until trainees can perform the entire and complex versions of the procedure independently. The Zwisch model can be helpful for teachers to gradually grant trainees more autonomy.

Focused training sessions
Although the integration of non-technical next to technical skills training is key, there is also room for focused training. When surgical procedures require skills that need to be automated (e.g., suturing) or require deeper understanding (e.g., models for closed loop communication), supplementary focused training sessions are installed with the specific goal to solely focus on, train, and acquire specific skills. If skills like suturing or closed loop communication are concurrently trained next to other technical and non-technical skills, it may cause cognitive overload. Then, teaching and training best occurs separately until the skills are partly or fully automated. As soon as these skills are acquired, trainees apply them in the integrated training sessions. By then, trainees are experienced on that part of the procedure and have more cognitive resources available to focus on other important technical and non-technical skills. Focused training sessions typically occur in a simulated setting, with or without a teacher being present, depending on the possibility for perceptual feedback. With some skills, after the teacher explained the skill by the Peyton four step approach, for example, the trainee can independently practice to automation with at-home training kits, simulators, etc. Figure 4 provides examples of training in simulated settings, both for focused training (A and B) and integrated training (C).

Individualized modules

When applying the aforementioned training and teaching principles, it is important to realize that there is no one size fits all approach. In close and continuous consultation, surgical teachers and trainees individualize each training module and the training sessions within each training module, according to each individual trainee's experience, progress, needs, or interests. Skills and abilities that are already obtained move to the background to enable focus on other skills and abilities. The length of each training module is flexible and different per individual trainee depending on the trainee's speed of progress. Different training modules do not necessarily succeed each other but rather run parallel or partly parallel if similarities in difficulty, techniques, and trainee requirements allow. The basic principle is and remains the training of entire surgical procedures in which both technical and non-technical skills have to be applied, starting simple and gradually move on to a highly complex endpoint. Along this process, ideally, trainees are guided by two or maximum three teachers within a module; that way, teachers get to know their trainees' strengths and weaknesses, which helps the teacher to tailor the teaching, but also offers trainees feedback from different perspectives.

Within training sessions

Good teaching during training sessions is important to develop trainees. Teaching best occurs supportive and in a safe environment. This requires teachers to take time and be calm, approachable, and respectful. Good teaching contains instructions, explanations, demonstrations, and honest feedback with suggestions for improvement tailored to each individual trainee's needs and level. Effective feedback is constructive, non-offensive, and neutrally formulated. The trainee's strong points and weaknesses are addressed based on and illustrated with first-hand observations. The goal of feedback is to improve or maintain the trainee's level of performance in the future.

The BID model can be helpful to structure the teaching within training sessions. During the briefing, the teacher and trainees briefly discuss and agree on personal learning goals for the upcoming training session, the guidance expectations, and the training activities. Guidance can range from strict and continuously to distant and minimal. Training activities can vary from observing to performing the entire procedure. In integrated training sessions, the teacher and trainee work on one technical and one personal non-technical learning goal. In focused training sessions, the teacher and trainee work on either one or two technical or one or two non-technical personal learning goals.

During the hands-on training phase, the actual teaching and training take place. How the teacher teaches depends on the nature of the training session (integrated or focused) and the trainees' experience. Observational learning, a high level of teaching intensity, and a low level of trainee contribution are typical for inexperienced trainees. Distant monitoring, a low level of teaching intensity, and a high level of trainee contribution are typical for experienced trainees. The Zwisch model can help teachers to grant autonomy and determine the teaching intensity.

Figure 4. Three simulated settings.
Three simulated training settings, ranked from least realistic and most controlled (A) to most realistic and least controlled (C). (A) A focused training session regarding technical laparoscopic skills using a virtual reality simulator. Basic laparoscopic skills are automated before trainees apply them in the operating room. (B) A focused training session regarding non-technical communication and team skills using a human patient simulator with video recordings for debriefing (view from control room). Additional training may occur next to training in the real-life operating room. (C) An integrated training session using an animal model. Such training is scheduled best if trainees have gained sufficient experience regarding the necessary technical and non-technical skills.

During technically focused training sessions, trainees are preferably only taught how to perform skills technically (by the Peyton four step approach, for example). Teaching is predominantly directive and instructive and trainees' mistakes are corrected immediately. During nontechnically focused training sessions, teaching may occur more reflective, in debriefings or pause and reflect procedures. If, for example, a communication model is new to trainees, they may first receive short instructions, then observe good and wrong examples, and then immediately apply the skill in a simulated setting. The trainees learn by the feedback provided and reflective questions asked by the teacher. Central questions may be: Why did it go the way it went (addressing both positive and corrective aspects)? Why did you act the way you acted? What are the (possible) consequences? What are the solutions for future practice? Video-recordings of the trainees' performance may be helpful.
In integrated training sessions, teaching focuses on both technical and non-technical skills concurrently. However, if some technical or non-technical skills are not yet sufficiently understood or trained, also within integrated training sessions, there can be episodes of pure technical and non-technical skills teaching. More focused training sessions may be required and can be installed according to the teacher's insight or on the trainee's request.
Each training session closes with a debriefing. A debriefing can typically start with the teacher asking the trainee what he or she thinks went well and needs to be improved, with a special focus on the trainee's personal technical and non-technical learning goals. The teacher has observed and analyzed the trainee's performance, preferably by using OTAS or NOTSS (non-technical skills) and OSATS (technical skills). Observations may focus on a few points, adjusted to the trainees' personal learning goals. The teacher provides feedback in a dialog with the trainee and tries to get thoughts, beliefs, and reasoning processes clear. The teacher supports the trainee in his or her development by adding knowledge and expertise to the discussion. The briefing finishes with the trainee formulating future personal learning goals or intentions for the next training.

REFERENCES

1. Hull L, Sevdalis N. Advances in teaching and assessing nontechnical skills. Surgical Clinics, 2015. 95(4): p. 869-884.
2. Dedy NJ, et al. Teaching nontechnical skills in surgical residency: a systematic review of current approaches and outcomes. Surgery, 2013. 154(5): p. 1000-1008.
3. Dawe S, et al. Systematic review of skills transfer after surgical simulation-based training. British Journal of Surgery, 2014. 101(9): p. 1063-1076.
4. Whittaker G, et al. Teamwork assessment tools in modern surgical practice: a systematic review. Surgery research and practice, 2015. 2015.
5. Brunckhorst O, et al. The relationship between technical and nontechnical skills within a simulation-based ureteroscopy training environment. Journal of surgical education, 2015. 72(5): p. 1039-1044.
6. Agha RA, Fowler AJ, Sevdalis N. The role of non-technical skills in surgery. Annals of Medicine and Surgery, 2015. 4(4): p. 422-427.
7. Gjeraa K, et al. Non-technical skills in minimally invasive surgery teams: a systematic review. Surgical endoscopy, 2016. 30(12): p. 5185-5199.
8. Hull L, et al. Observational Teamwork Assessment for Surgery: Content Validation and Tool Refinement. Journal of the American College of Surgeons, 2011. 212(2): p. 234-243. e5.
9. Yule S, et al. Surgeons' Non-technical Skills in the Operating Room: Reliability Testing of the NOTSS Behavior Rating System. World Journal of Surgery, 2008. 32(4): p. 548-556.
10. White C, Rodger MW, Tang T. Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. The Obstetrician & Gynaecologist, 2016. 18(1): p. 53-63.
11. Nicholls D, et al. Teaching psychomotor skills in the twenty-first century: revisiting and reviewing instructional approaches through the lens of contemporary literature. Medical teacher, 2016. 38(10): p. 1056-1063.
12. LeBlanc VR, et al. 18 Simulation in Postgraduate Medical Education. 2011.
13. McKendy KM, et al. Perioperative feedback in surgical training: A systematic review. Am J Surg, 2017. 214(1): p. 117-126.
14. Timberlake MD, et al. What do we know about intraoperative teaching?: A systematic review. Annals of surgery, 2017. 266(2): p. 251-259.
15. Spanager L, et al. Comprehensive feedback on trainee surgeons' non-technical skills. International journal of medical education, 2015. 6: p. 4.
16. Alken A, et al. Teaching complex surgical psychomotor skills in the OR: a BEME systematic review. Medical teacher, Under review.
17. Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance patient safety: a systematic review. Medical education, 2012. 46(11): p. 1042-1054.
18. Kneebone R, et al. Distributed simulation--accessible immersive training. Med Teach, 2010. 32(1): p. 65-70.
19. Ahmed M, et al. Actual vs perceived performance debriefing in surgery: practice far from perfect. The American Journal of Surgery, 2013. 205(4): p. 434-440.

20. Palter VN, Grantcharov TP. Simulation in surgical education. Canadian Medical Association Journal, 2010. 182(11): p. 1191-1196.
21. Sadideen H, et al. Simulators and the simulation environment: Getting the balance right in simulation-based surgical education. International Journal of Surgery, 2012. 10(9): p. 458-462.
22. Alken A, et al. Integrating technical and non-technical skills coaching in an acute trauma surgery team training: Is it too much? The American Journal of Surgery, 2017.
23. Alken A, et al. Feedback activities of instructors during a trauma surgery course. The American Journal of Surgery, 2013. 206(4): p. 599-604.
24. DaRosa DA, et al. A Theory-Based Model for Teaching and Assessing Residents in the Operating Room. Journal of Surgical Education, 2013. 70(1): p. 24-30.
25. Roberts NK, et al. The Briefing, Intraoperative Teaching, Debriefing Model for Teaching in the Operating Room. Journal of the American College of Surgeons. 208(2): p. 299-303.
26. Meyerson SL, et al. Defining the Autonomy Gap: When Expectations Do Not Meet Reality in the Operating Room. Journal of Surgical Education. 71(6): p. e64-e72.
27. Van Merriënboer JJ, Kirschner PA, Kester L. Taking the load off a learner's mind: Instructional design for complex learning. Educational psychologist, 2003. 38(1): p. 5-13.
28. Janssen-Noordman, A.M., et al., Design of integrated practice for learning professional competences. Med Teach, 2006. 28(5): p. 447-52.

8

GENERAL DISCUSSION WITH THOUGHTS FOR FUTURE RESEARCH

To date, most research to effective coaching has been focusing on the effects of conceptual coaching frameworks rather than specific coaching behaviors.[1,2] This makes it hard to draw conclusions on *what* coaching behaviors surgical teachers should apply *in what way* to effectively teach technical and non-technical skills in hands-on surgical training. Future research should be focused on investigating the effectiveness of specific coaching behaviors. Investigating and determining what specific coaching behaviors are effective will be challenging. Research to effective feedback in clinical training suggests that it is not only important how teachers construct and deliver feedback, but that it is just as important how their trainees receive and interpret the provided feedback.[3] The effectiveness of coaching seems to depend on the interplay between what coaching behaviors are applied and how they are experienced by the trainees. Research should combine objective analyses with subjective perceptions of specific coaching behaviors.

Currently, the effectiveness of surgical coaching is commonly analyzed through trainee and teacher perception questionnaires.[4] Such questionnaires provide insight in how trainees and teachers experienced the coaching. However, they appear to be not very reliable since perception questionnaires have been shown to differ substantially from independent observations[4,5] and to be prone to recall and conformation bias.[4] Furthermore, they only provide limited insight in how specific coaching behaviors are interpreted and experienced by trainees. Several researchers have developed observational instruments to enable a more objective way to analyze surgical coaching behaviors.[6-13] Most instruments provide insight in what behaviors are applied by teachers, but not in the effectiveness or quality of these behaviors.[6-12] To our knowledge, only one instrument, the Structured Training Trainer Assessment Report (STTAR), enables an assessment of the quality of the observed coaching behaviors.[13] The STTAR is an observational instrument which enables independent observers to elaborately analyze what coaching behaviors surgical teachers apply and to rate the quality of this behavior in terms of content and the way it is applied. The aim of the STTAR is to provide surgical teachers with an independent analysis of their coaching and define areas for improvement. Since the STTAR is a complex and labor-intensive instrument the researchers also developed a less labor-intensive alternative: the mini-STTAR. This mini-STTAR is a web-based perception questionnaire for trainees to assess the quality of coaching and provide feedback to their teachers.[13] Although designed as two separate entities, to assess the effectiveness of coaching, it may be more effective to combine the objective observations of the STTAR with the subjective perceptions of the mini-STTAR. Further research is needed to investigate the possibilities for and effectiveness of such a combination to analyze surgical teachers' coaching quality.

Recently, a mixed method approach has been proposed to analyze the effectiveness of surgical coaching.[4] Video recordings made during hands-on surgical training are reviewed by teachers and trainees, complemented with transcriptions of the conversations, which enables face-to-face discussions regarding the perceptions and interpretations of specific coaching behaviors. In this way objective analyses are combined with subjective perceptions. By simultaneous involving different stakeholders in the reviewing process it is possible to investigate what exact coaching behaviors teachers applied with what intentions, discuss with trainees whether this was effective or not, why this was effective or not, and what alternatives are. Education specialists can be involved in these discussions. Objective surgical skills assessments may also be involved to gain insight in the objective effects of coaching behaviors on trainees' skill acquisition. Measurements of emotional status of the teacher and trainee can also be added to explore the relation between stress, stress coping, applied coaching behaviors and learning by the trainees. In our research we started to explore the role of stress by including time-related stress measurements recorded during surgical hands-on training with the Vital Connect HealthPatch™ (Campbell, CA, USA).[14]
The proposed mixed method approach also enables trainees to provide their teachers with detailed feedback on their coaching behavior. The effect of face-to-face feedback sessions based on video review has not yet been investigated. Possibly, this is a more reliable and effective method to gain insight in trainees' interpretations and experiences of specific coaching behaviors than perception questionnaires. A main drawback of the mixed method approach appears to be the large time investment necessary to generate feedback. Other methods in which trainees provide face-to-face feedback to their teachers appear less time consuming and show promising results in improving teachers' clinical hands-on coaching.[15]

As mentioned earlier, the effectiveness of coaching in surgical hands-on training seems to depend on the interplay between the content of the coaching and the way it is provided by the teachers on the one hand, and the interpretation and experience of the coaching by the trainees on the other hand. Conducting research to this interplay will be highly challenging since many variables can be involved. Research to clinical training suggests that the effectiveness of feedback depends on:

- the task characteristics (e.g. nature and complexity);
- the trainees and teachers who are involved (e.g. prior experiences, expectations, personalities and relations);
- the trainees' performance (e.g. below or above expectations);
- the content of the message (e.g. intensity and complexity) and the way it is provided (e.g. respectful, related to trainees' learning goals);
- the context in which the message is provided (e.g. stress);
- the culture in which the message is provided (e.g. common implicit practice).

How these variables exactly relate to the effectiveness of feedback, how variables interact with one another, and what ultimately determines whether a feedback message contributes to or compromises the development of trainees, is vastly unknown.[3,16] The effectiveness of feedback appears to be hard to predict[16] and to be variable.[3] We expect that the same is true for surgical coaching behavior. Further research is needed on what variables are involved in training situations, how these variables interact and what ultimately determines whether surgical coaching behavior is effective or not. Given all the variables involved, this will be a highly complex undertaking. We advise to start researching the reception and interpretation of coaching behaviors by trainees in relation to the intentions of teachers and objective analyses of the coaching behaviors that were applied in circumstances and training situations where different variables can be controlled (as outlined at previous page). This may lead to a new educational framework for effective hands-on coaching in surgery. We, however, realize that the usability of such a framework may be limited to certain surgical teaching situations, contexts, cultures and even personalities.

REFERENCES

1. Alken AL, Jan-Maarten; Thornblade, Lucas; Chen, Xiaodong; Hull, Louise; Horvath, Karen; van Goor, Harry; Fluit, Cornelia, Evidence-based effective teaching behaviors for complex psychomotor skills training. Submitted.
2. Huang GC, et al. Procedural instruction in invasive bedside procedures: a systematic review and meta-analysis of effective teaching approaches. BMJ Qual Saf, 2016. 25(4): p. 281-94.
3. Lefroy J, et al. Guidelines: the do's, don'ts and don't knows of feedback for clinical education. Perspect Med Educ, 2015. 4(6): p. 284-99.
4. Mazer LM, et al. Evaluating Surgical Coaching: A Mixed Methods Approach Reveals More Than Surveys Alone. J Surg Educ, 2018.
5. Alken A, et al. Coaching during a trauma surgery team training: perceptions versus structured observations. Am J Surg, 2015. 209(1): p. 163-9.
6. Blom EM, et al. Analysis of verbal communication during teaching in the operating room and the potentials for surgical training. Surg Endosc, 2007. 21(9): p. 1560-6.
7. Chen XP, et al. A taxonomy of surgeons' guiding behaviors in the operating room. Am J Surg, 2015. 209(1): p. 15-20.
8. Hauge LS, Wanzek JA, Godellas C. The reliability of an instrument for identifying and quantifying surgeons' teaching in the operating room. Am J Surg, 2001. 181(4): p. 333-7.
9. Roberts NK, et al. Capturing the teachable moment: a grounded theory study of verbal teaching interactions in the operating room. Surgery, 2012. 151(5): p. 643-50.
10. Sutkin G, Littleton EB, Kanter SL. How surgical mentors teach: a classification of in vivo teaching behaviors part 1: verbal teaching guidance. J Surg Educ, 2015. 72(2): p. 243-50.
11. Alken A, et al. Integrating technical and non-technical skills coaching in an acute trauma surgery team training: Is it too much? The American Journal of Surgery, 2017.
12. Alken A, et al. Feedback activities of instructors during a trauma surgery course. The American Journal of Surgery, 2013. 206(4): p. 599-604.
13. Wyles SM, et al. Development and implementation of the Structured Training Trainer Assessment Report (STTAR) in the English National Training Programme for laparoscopic colorectal surgery. Surg Endosc, 2016. 30(3): p. 993-1003.
14. Alken A, et al. Measuring stress and coaching behaviors during a highly realistic trauma surgery team training. Submitted.
15. Fluit CR, et al. Repeated evaluations of the quality of clinical teaching by residents. Perspectives on medical education, 2013. 2(2): p. 87-94.
16. van de Ridder J, et al. Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. Medical education, 2015. 49(7): p. 658-673.

8

CHAPTER 9

Summary
Samenvatting

SUMMARY

Effective surgery requires high-quality technical skills and non-technical communication and team skills. By tradition, surgical training has a strong focus on the teaching of technical skills. Recently the attention for teaching non-technical skills increased, however, it remains a minimal, infrequent and unstructured part of surgical training.

Important for the effectiveness of surgical training is the coaching behavior that surgical teachers apply. The specific content of this coaching behavior and how teachers balance technical and non-technical skills teaching during training is largely unknown.

Most surgical training occurs in the operating room. Under the supervision of experienced surgeons surgical trainees train and perform surgery on real patients. Training in the operating room is highly complex since the environment is high-stake, variable and stressful, and the skills that have to be trained are challenging. This environment may be prone to disruptive communication and coaching behavior with negative consequences for the quality of training.

By the use of simulation, surgical skills can also be taught and trained outside the operating room. Simulation ranges from simple bench top models to highly realistic live animal models. Different simulation models have different consequences regarding training time, teaching strategies and the reality of the training environment. When live animal models are used, such a training environment neatly represents the hectic real-life operating room.

The objectives of this thesis were:

1. to categorize and analyze the coaching behaviors surgical teachers apply when teaching in the operating room;
2. to determine how these teaching behaviors are perceived by surgical teachers and trainees, and how they relate to independent observations;
3. to investigate if the 'priming' of teachers results in an increase in non-technical skills coaching;
4. to determine stress of surgeons and residents in daily surgical practice;
5. to determine how surgical teachers experience stress during teaching in the operating room and how this relates to their coaching behaviors;
6. to review evidence-based coaching behaviors that are effective in teaching complex psychomotor skills;
7. to provide an overview on how to integrate technical and non-technical skills teaching and training in the operating room.

In four studies (chapter two to five) we collected data during a highly realistic acute trauma surgery team training, which was part of the DSATC course combining the Definitive Surgical (DSTC) and Definitive Anesthetic Trauma Care (DATC) courses. In this team training, twelve operating teams in simulated operating rooms dealt with six consecutive life-threatening

9

trauma scenarios for which injuries were inflicted on anesthetized live porcine models. In each team, one certified and experienced surgical teacher coached two surgical trainees while operating with an anesthetic team and a scrub nurse. Surgical trainees had to gain control of the "patient" by applying the necessary technical and non-technical skills.

Chapter One describes the general introduction of this thesis and provides the thesis outline.
In **Chapter Two** we aimed to analyze the feedback and instruction behaviors that surgical teachers apply during the acute trauma surgery team training. Data of eleven surgical teachers were collected. All operating activities were audio and videotaped. For each teacher forty-five minutes of identical operating scenarios were reviewed and analyzed with a newly designed observational coaching instrument. The instrument distinguishes feedback and instruction, with different levels of specific information, directed on technical or non-technical team and communication skills.
The surgical teachers provided instruction significantly more often than feedback. Most feedback and instruction contained nonspecific or low specific information and were directed on technical skills. The coaching on non-technical communication skills contained more specific information than the coaching on other skills. We concluded that surgical teachers predominantly focus on the coaching of technical skills with unspecific instructions. The observational coaching instrument enabled the scoring of different coaching activities undertaken by teachers.
We analyzed the perceptions of the surgical trainees and teachers regarding the quality of the coaching provided during the acute trauma surgery team training, and compared these to structured observations of the applied coaching behaviors in **Chapter Three**. For each teacher, forty-five minutes of identical operating scenarios were analyzed with the observational instrument to quantitatively asses the use of positive feedback, corrective feedback, instruction, and how and why explanations. The observed coaching behaviors were then compared to the perceptions of these behaviors as reported on questionnaires by the trainees and teachers.
The ratings regarding the perceived coaching quality differed between trainees and teachers, with trainees generally giving higher ratings. While both the trainees and the teachers gave high ratings for non-technical skills coaching, for the provision of corrective and complimenting feedback and for explanations why, the structured observations scored lower for these categories. In conclusion, trainees and teachers both overrate coaching. Trainee questionnaires and teacher self-reports based on subjective perceptions neither reflect the actual coaching activities, nor identify coaching skill deficits.
The aim of **Chapter Four** was to investigate whether an additional instruction regarding non-technical skills before the acute trauma surgery team training would increase the frequency of non-technical skills coaching by surgical teachers during the training. Twelve surgical teachers participated. Before the start of the training, six teachers were 'primed' by receiving additional instruction on non-technical skills coaching. The remaining six received no additional

instruction. All coaching activities were recorded on audio and video. For each teacher, two blinded observers scored for two of the six trauma scenarios recorded whether the applied coaching was directed on technical or non-technical skills. The frequency of the non-technical skills coaching was compared between the 'primed' and the 'non-primed' teachers, also taking into account differences in trainees' level of experience.

The application of non-technical skills coaching during the training was very low without a difference between the 'primed' teachers who received additional instruction and those who did not (the 'non-primed' teachers). We found no effects according to different trainee levels. A short additional instruction regarding coaching on non-technical skills before training does not increase the frequency of coaching on these skills by experienced trainers, possibly due to the complexity of the acute trauma surgery team training course.

Chapter Five describes a study in which we explored the stress patterns in consultant, fellow and resident surgeons during their daily activities at work, including operating room, outpatient clinic, ward and administrative activities. An additional reason for this study was to gain experience regarding the feasibility, usability and reliability of real-life continuous stress measurements with an innovative wearable device in preparation for a consequent study relating stress to coaching behavior (see Chapter six). For two to three consecutive days participants were attached to a small, light-weighted adhesive patch, the Vital Connect HealthPatch™ (the current version is called VitalPatch®, Campbell, CA, USA). This patch measures heart rate (HR), heart rate variability (HRV) and calculates a real-time stress percentage according to a validated algorithm based on HR and HRV. All participants kept a logbook in which they noted the time and activities they had undertaken. We compared the objective HRV data and stress percentage with baseline values and between activities in consultants, fellows/senior residents, and junior residents. We also compared the HRV data and stress percentage collected during surgery in the operating room with the subjective self-perceived stress scores as reported by the participants on the State Trait Anxiety Inventory (STAI).

In twenty participants (five consultants, seven fellows/senior residents, eight junior residents) the HRV data and stress percentage indicated significantly higher stress during operating than during other daily activities. Consultant surgeons in comparison to fellows/senior residents showed lower stress while operating. Objective stress as measured by the HealthPatch™ did not correlate with subjective stress as measured by the STAI during operating theatre activities. Continuous stress monitoring using a wearable sensor is feasible and reveals relevant differences in stress according to type of daily activity and surgical experience. Lack of correlation with subjective stress is possibly due to biased stress reporting.

The objective in **Chapter Six** was to investigate the relation between stress in surgical teachers and their coaching behaviors, regarding technical and non-technical skills, during the acute trauma surgery team training. A second objective was to assess the minute-to-minute relationship between salient events during training and stress levels of a teacher and trainee working in the same team. Nine teachers and one trainee were equipped with the Vital Connect

HealthPatch™. The Low Frequency / High Frequency (LF/HF) ratio was used as a HRV proxy for stress. All coaching activities were recorded on audio and video. For each teacher, four 10-minute time frames were determined representing the two lowest and two highest mean LF/HF-ratios. A blinded observer categorized and calculated the frequency of all technical and non-technical coaching behaviors. The relationship between salient events and stress was explored by comparing two 30-minute time periods containing steep increases and decreases in the LF/HF-ratio of one teacher and one trainee in the same team, and the occurrence of salient events collected by a blind observer.

Usable data was extracted for five teachers and one trainee. The teachers' coaching behaviors did not differ between periods with the highest and lowest mean LF/HF-ratios. Steep increases in the LF/HF-ratio of the trainee coincided with the infliction of an injury. In conclusion, changes in stress level do not correlate with changes in coaching behavior regarding technical and non-technical skills of experienced trainers in an acute trauma surgery simulation training. Trainee stress appears to be related to the infliction of life threatening injuries that have to be controlled.

Chapter Seven contains a literature review study that we conducted to identify evidence-based effective teaching behaviors that increased complex psychomotor skills of trainees as measured by objective means. We searched the MEDLINE, PsycINFO and ERIC databases for relevant papers on surgical, medical, sports and music training. This multifaceted approach was chosen because these areas of training share similar psychomotor skills characteristics and complexity. We identified teaching behaviors and assessed the impact levels and quality of the included studies with the Medical Education Research Study Quality Instrument (MERSQI) and the Newcastle-Ottawa Scale-Education (NOS-E).

Of 18,337 references seven studies, of which one concerned surgical training, met the inclusion criteria. The teaching behaviors that showed to objectively improve skills in trainees are feedback, instruction, active trainee involvement and demonstration. Feedback and instruction with an external focus, stimulating the trainee to focus on the task and effect rather than on how to move, was supported by the strongest evidence. Harshly criticizing and belittling teaching behavior was related to adverse effects on trainees' skills acquisition. The data quality and evidence for most of the identified teaching behaviors were weak with low impact levels. In conclusion, feedback, instruction, active trainee involvement and demonstration seem to improve the development of complex psychomotor skills in hands-on training. However, evidence supporting the effectiveness is generally weak. More and better quality research is needed which assesses the effects of teaching behaviors on the acquisition of complex surgical skills in trainees, as defined by objective outcome measures.

In **Chapter Eight** we summarize the state of the art knowledge on the teaching and training of technical and non-technical skills in hands-on surgical training with varying complexity, partly based on the research we have conducted. We propose a theoretical framework which can be helpful for integrating technical and non-technical skills coaching in surgical training.

We end with a short general discussion and thoughts for future research on effective surgical coaching.

SAMENVATTING

Om als chirurg te kunnen werken moet er een specialistische opleiding worden gevolgd. Tijdens deze opleiding worden chirurgen in opleiding getraind door chirurgische docenten. Van oorsprong zijn chirurgische trainingen gericht op het ontwikkelen van technische vaardigheden. Echter, om effectief en veilig te kunnen opereren zijn niet alleen goed ontwikkelde technische vaardigheden noodzakelijk, maar ook goed ontwikkelde niet-technische vaardigheden, zoals communicatie- en teamvaardigheden. Chirurgen werken immers vrijwel altijd in teamverband. Het belang van het ontwikkelen van deze niet-technische vaardigheden wordt pas sinds kort onderkend. Het expliciet aanleren van niet-technische vaardigheden in chirurgische trainingen vindt dan ook nog maar mondjesmaat, en meestal op ongestructureerde wijze plaats.

Het trainen van chirurgen in opleiding vindt hoofdzakelijk plaats in de operatiekamer. Tijdens operaties op 'echte' patiënten oefenen en ontwikkelen chirurgen in opleiding hun vaardigheden onder begeleiding van ervaren chirurgische docenten. De chirurgische docent speelt een belangrijke rol in de kwaliteit van deze trainingen. Zijn of haar coachingsgedrag bepaalt voor een belangrijk deel hoe effectief de training is. Er is nog maar weinig bekend over hoe dit coachingsgedrag er inhoudelijk uitziet en hoe docenten daarbij hun aandacht verdelen over technische en niet-technische vaardigheden.

De operatiekamer is een zeer complexe omgeving om in te trainen. Niet alleen zijn de vaardigheden die geoefend worden op zichzelf ingewikkeld. De omstandigheden waaronder geoefend wordt zijn ook nog eens variabel en stressvol. Tegelijkertijd dient de patiënt de beste en meest veilige zorg te krijgen. Dergelijke complexe omstandigheden kunnen het coachingsgedrag van chirurgische docenten negatief beïnvloeden en de kwaliteit van deze trainingen schaden. Door gebruik te maken van simulatie zijn er mogelijkheden om in een minder complexe omgeving te trainen, buiten de operatiekamer. Chirurgen in opleiding kunnen hun vaardigheden oefenen en ontwikkelen op simulatiemodellen, variërend van eenvoudige (kunststof) modellen tot levende diermodellen die de realiteit van de hectische operatiekamer zeer dicht benaderen. De diverse simulatiemodellen hebben verschillende voor- en nadelen met betrekking tot de tijd die genomen kan worden om te trainen, de mogelijke onderwijsstrategieën en de natuurgetrouwe benadering van de trainingsomgeving.

De studies in dit proefschrift zijn verricht met de volgende doelstellingen:

1. categoriseren en analyseren van coachingsgedrag dat chirurgische docenten toepassen in de operatiekamer tijdens het trainen van chirurgen in opleiding;
2. onderzoeken hoe dit coachingsgedrag wordt beoordeeld door de docenten zelf en door de chirurgen in opleiding, en hoe deze beoordelingen zich verhouden tot observaties van dit coachingsgedrag door derden;

3. onderzoeken of het 'primen' (kort instrueren) van chirurgische docenten voorafgaand aan een training leidt tot een toename van de coaching op niet-technische vaardigheden;
4. onderzoeken in hoeverre chirurgen en chirurgen in opleiding stress ervaren tijdens hun dagelijkse werkzaamheden, en in de operatiekamer in het bijzonder;
5. onderzoeken in hoeverre chirurgische docenten stress ervaren tijdens het trainen van chirurgen in opleiding in de operatiekamer, en of er een relatie is tussen stress en het toegepaste coachingsgedrag;
6. systematisch onderzoeken welke coachingsgedragingen effectief zijn in het aanleren van complexe psychomotorische vaardigheden;
7. aanbevelingen formuleren met betrekking tot de integratie van technische en niet-technische vaardigheidstraining tijdens chirurgische trainingen in de operatiekamer.

In dit proefschrift zijn zes studies uitgevoerd. Voor vier studies (hoofdstuk twee tot en met vijf) hebben we gegevens verzameld tijdens de teamtraining acute traumachirurgie, die onderdeel is van de gecombineerde Definitive Surgical Trauma Care (DSTC) en Definitive Anaesthetic Trauma Care (DATC) cursus: de Definitive Surgical and Anaesthetic Trauma Care (DSATC) cursus. Deze teamtraining heeft als doel om chirurgen in opleiding de technische en niet-technische communicatie- en teamvaardigheden aan te leren die nodig zijn om een levensbedreigde 'patiënt' te behandelen tijdens een spoedoperatie. Tijdens deze hoogrealistische teamtraining oefenen twaalf operatieteams zes levensbedreigende traumascenario's op varkens onder narcose in een gesimuleerde operatiekamer. Elk traumascenario betreft een specifieke (set van) verwonding(en) waarvoor verschillende chirurgisch-technische verrichtingen nodig zijn. Binnen elk operatieteam voeren twee chirurgen in opleiding de operaties uit, in samenwerking met een anesthesieteam en een operatieassistent. De twee chirurgen in opleiding worden daarbij gecoacht door één gecertificeerde en ervaren traumachirurgische docent.

Hoofdstuk Een van dit proefschrift voorziet in een algemene introductie en beschrijft de opzet van het proefschrift.
In **Hoofdstuk Twee** hebben we het feedback- en instructiegedrag van traumachirurgische docenten in kaart gebracht waarmee zij chirurgen in opleiding trainen tijdens de teamtraining acute traumachirurgie. We verzamelden daarvoor onderzoeksgegevens bij elf chirurgische docenten. Van alle activiteiten die werden uitgevoerd tijdens de traumascenario's van deze teamtraining werden audio- en video-opnamen gemaakt. Vervolgens werd met behulp van een nieuw ontworpen observatie-instrument van elke chirurgische docent dezelfde vijfenveertig minuten aan scenario's systematisch geanalyseerd. Met het observatie-instrument kon onderscheid worden gemaakt tussen feedback en instructie, tussen de verschillende niveaus van specifieke informatie die werd gegeven, en tussen het aanleren van technische vaardigheden, teamvaardigheden of communicatievaardigheden. Chirurgische docenten

bleken significant vaker instructie te geven dan feedback. De meeste feedback en instructies gingen gepaard met niet- of weinig specifieke informatie, en waren gericht op het aanleren van technische vaardigheden. De coaching die gericht was op communicatievaardigheden bevatte specifiekere informatie dan de coaching die gericht was op andere vaardigheden. We stelden vast dat chirurgische docenten tijdens de teamtraining acute traumachirurgie voornamelijk op technische vaardigheden coachten, en door middel van niet- of weinig specifieke instructies. Het observatie-instrument bleek geschikt om de verschillende coachingactiviteiten van chirurgische docenten te onderscheiden.

We onderzochten ook hoe de chirurgische docenten en de chirurgen in opleiding de kwaliteit (effectiviteit, leerzaamheid, etc.) van de coaching tijdens de teamtraining acute traumachirurgie ervoeren. We vergeleken de ervaringen van beide groepen met elkaar, en met gestructureerde observaties van het coachingsgedrag. Deze studie staat beschreven in **Hoofdstuk Drie**. Met behulp van een observatie-instrument analyseerden we van elke chirurgische docent het coachingsgedrag binnen dezelfde vijfenveertig minuten van de traumascenario's. Daarbij werd de hoeveelheid positieve feedback, corrigerende feedback, instructie en specifieke informatie (met uitleg over het hoe en waarom) in kaart gebracht. Vervolgens werden deze bevindingen vergeleken met hoe de chirurgische docenten en de chirurgen in opleiding het coachingsgedrag hadden beleefd en gescoord op vragenlijsten. De ervaringen van de chirurgische docenten en chirurgen in opleiding verschilden: de chirurgen in opleiding gaven in het algemeen hogere scores aan de kwaliteit van coaching dan de chirurgische docenten. Beide groepen gaven hoge scores aan de coaching op niet-technische vaardigheden, het toepassen van corrigerende en complimenterende feedback, en het geven van uitleg waarom, terwijl dit nauwelijks was waargenomen tijdens de gestructureerde observaties. Vergeleken met deze gestructureerde observaties bleken de chirurgische docenten en de chirurgen in opleiding de kwaliteit van de coaching te onderschatten. Subjectieve vragenlijsten waren niet goed in staat om tekortkomingen in het coachingsgedrag van docenten vast te stellen tijdens een teamtraining acute traumachirurgie.

In **Hoofdstuk Vier** hebben we onderzocht of een aanvullende instructie aan chirurgische docenten tot meer coaching op niet-technische team- en communicatievaardigheden leidt tijdens de teamtraining acute traumachirurgie. Twaalf chirurgische docenten participeerden in dit onderzoek. Voor aanvang van de teamtraining ontvingen zes van hen een korte, aanvullende instructie ('priming'), specifiek gericht op het coachen van niet-technische vaardigheden ('geprimede' docenten). De overige zes docenten ontvingen geen aanvullende instructies ('niet-geprimede' docenten). Tijdens de teamtraining werden alle activiteiten vastgelegd op audio en video. Van elke docent werden dezelfde twee (van in totaal zes) traumascenario's geobserveerd door twee observatoren die niet wisten of zij met 'geprimede' of 'niet-geprimede' docenten te maken hadden. De observatoren analyseerden of de coaching van de docenten gericht was op technische of niet-technische vaardigheden. Vervolgens werd de frequentie van de coaching op niet-technische vaardigheden tussen de 'geprimede' en 'niet-geprimede'

9

docenten met elkaar vergeleken. Daarbij werd ook onderzocht of er een relatie was met de verschillende ervaringsniveaus van de chirurgen in opleiding die werden gecoacht. Er was geen verschil tussen beide groepen docenten wat betreft de hoeveelheid coaching op niet-technische vaardigheden. In zowel de 'geprimede' als de 'niet-geprimede' groep docenten was de coaching op niet-technische vaardigheden erg laag. Ook waren er waren geen verschillen in relatie tot de ervaringsniveaus van de chirurgen in opleiding. We concludeerden dat een korte, aanvullende instructie aan ervaren chirurgische docenten gericht op niet-technische vaardigheden voorafgaand aan een teamtraining acute traumachirurgie, niet leidt tot meer coaching op deze vaardigheden. De complexiteit van deze teamtraining verklaart mogelijk het uitblijven van een verschil.

In **Hoofdstuk Vijf** hebben we de (mentale) stresspatronen van chirurgen met verschillende ervaringsniveaus tijdens hun dagelijkse werkzaamheden onderzocht met behulp van een nieuwe draagbare sensor: de Vital Connect HealthPatch™ (de huidige VitalPatch®). Deze HealthPatch™ is een kleine, lichtgewicht pleister, uitgerust met een sensor die continue metingen verricht van vitale parameters. Deze sensor registreert onder andere de hartslag (HS) en het elektrocardiogram (ECG), waaruit de hartslagvariabiliteit (HSV; als indicator voor stress) en een stresspercentage (gebaseerd op een algoritme uitgaande van de HS en HSV) kan worden berekend. Twintig deelnemers, waaronder vijf chirurgen, zeven fellow-chirurgen/ouderejaars assistenten in opleiding tot specialist (AIOS) en acht jongerejaars AIOS droegen deze HealthPatch™ gedurende twee tot drie dagen. De HSV en het stresspercentage werden geanalyseerd en vergeleken tijdens rust (baseline) en de verschillende dagelijkse activiteiten van chirurgen (opereren, patiënten behandelen op de polikliniek, patiënten behandelen op de verpleegafdeling en administratieve taken). Alle deelnemers noteerden hun dagelijkse activiteiten in een dagboek en er was dagelijks een 'debriefing'. Rondom het opereren werd ook de subjectieve stress van chirurgen vastgelegd middels de State Trait Anxiety Inventory (STAI). De HSV-gegevens en stresspercentages van de deelnemers toonden meer stress tijdens opereren dan tijdens de baselinemeting en de andere dagelijkse bezigheden. Ervaren chirurgen hadden minder stress tijdens het opereren dan fellow-chirurgen en ouderejaars AIOS. De objectieve stress, gemeten door deze HeathPatch™, correleerde niet met de subjectief gemeten stress op de STAI rondom de operaties. We concludeerden dat de HealthPatch™ goed bruikbaar is voor het continu meten van stress bij chirurgen (in opleiding). De sensor geeft goed inzicht in de mate van stress van chirurgen tijdens hun dagelijkse bezigheden en de individuele verschillen hierin. Het ontbreken van een correlatie tussen objectieve en subjectieve stress zou verklaard kunnen worden door een onderrapportage van stress door chirurgen.

In de studie beschreven in **Hoofdstuk Zes** onderzochten we de relatie tussen de stress van chirurgische docenten tijdens de teamtraining acute traumachirurgie, en het coachingsgedrag met betrekking tot technische en niet-technische vaardigheden. We verkenden ook de relatie tussen opvallende gebeurtenissen tijdens deze training (geagiteerde conversaties, aanbrengen van verwondingen, etc.) en de stressniveaus van één chirurgische docent en één chirurg in opleiding uit hetzelfde team. Negen docenten en één chirurg in opleiding werden uitgerust met de Vital Connect HealthPatch™ (de huidige VitalPatch®). Met deze sensor werd de Low Frequency/High Frequency (LF/HF-)ratio berekend als maat voor de HSV en als indicator voor stress. Alle coachingactiviteiten tijdens de training werden opgenomen op audio en video. Van elke docent werden vier periodes van 10 minuten training met de laagste twee en met de hoogste twee LF/HF-ratio gemiddelden geselecteerd. Op basis van de opnamen categoriseerde een geblindeerde onderzoeker (onwetend van het LF/HF-ratio gemiddelde) van elke docent alle coachingsgedragingen die te zien waren binnen deze periodes. De frequentie van alle technische en niet-technische coachingsgedragingen werd berekend. De relatie tussen opvallende gebeurtenissen en stress werd onderzocht door twee periodes van 30 minuten te selecteren waarin er sterke stijgingen en dalingen in de LF/HF-ratio van één docent en één chirurg in opleiding van hetzelfde team te zien waren. De tijdlijn van de stijgingen en dalingen werden vervolgens afgezet tegen de tijdlijn van de opvallende gebeurtenissen zoals vastgesteld door geblindeerde onderzoekers. Van vijf ervaren docenten en die ene chirurg in opleiding waren de stressdata bruikbaar voor analyse. Er werden geen verschillen gevonden in het coachingsgedrag van de docenten tijdens periodes van de hoogst gemiddelde LF/HF-ratio's in vergelijking met de periodes van de laagst gemiddelde LF/HF-ratio's. Bij de verkenning van de relatie tussen stress en opvallende gebeurtenissen werden steile stijgingen in de LF/HF-ratio van de chirurg in opleiding gezien op tijdstippen waarop de verwondingen werden aangebracht tijdens de gesimuleerde teamtraining acute traumachirurgie. Wij concludeerden dat veranderingen in het stressniveau van ervaren chirurgische docenten geen relatie lijken te hebben met veranderingen in het coachingsgedrag.

Hoofdstuk Zeven beschrijft een systematische literatuurstudie die wij uitvoerden naar effectieve coachingsgedragingen waarmee complexe, psychomotorische vaardigheden objectief aantoonbaar zijn aangeleerd bij mensen in opleiding tijdens 'hands-on' trainingssituaties. De databases MEDLINE, PsycINFO en ERIC werden doorzocht op relevante artikelen over het trainen van chirurgische, andere medische, sport- en muziekvaardigheden. Onderzoek naar het aanleren van sport- en muziekvaardigheden was onderdeel van deze studie omdat deze psychomotorische vaardigheden wat betreft trainingskarakter en complexiteit gelijkenissen vertonen met chirurgische psychomotorische vaardigheden. Studies met effectieve coachingsgedragingen werden geïdentificeerd en beoordeeld op kwaliteit en impactniveau met het Medical Education Research Study Quality Instrument (MERSQI) en de Newcastle-Ottawa Scale-Education (NOS-E). Van de totaal 18.337 referenties voldeden zeven studies aan onze inclusiecriteria. Slechts één studie daarvan ging over het trainen van chirurgische vaardigheden. Effectieve coachingsgedragingen betroffen feedback, instructie, het actief betrekken van mensen in opleiding tijdens de training, en het voordoen van vaardigheden. Het sterkste bewijs werd gevonden voor feedback en instructie met een externe focus, waarbij iemand zijn aandacht richt op de taak en het effect in plaats van hoe de beweging uitgevoerd moet worden. Coachingsgedrag waarbij mensen in opleiding buitensporig kritisch en kleinerend werden benaderd had negatieve effecten op het aanleren van vaardigheden. De kwaliteit van de studies en het bewijs voor de meeste van de, door ons geïdentificeerde coachingsgedragingen, waren zwak met een laag impactniveau. Feedback, instructie, het actief betrekken van mensen bij de training en het voordoen van vaardigheden tijdens 'hands-on' trainingssituaties hebben het grootste effect op het aanleren van complexe psychomotorische vaardigheden. Echter, het bewijs voor de effectiviteit van deze gedragingen is zwak. Verder onderzoek met objectieve uitkomstmaten is nodig om de effecten van het coachingsgedrag van docenten op het aanleren van complexe chirurgische vaardigheden bij chirurgen in opleiding te onderzoeken.

In **Hoofdstuk Acht** geven we, mede op basis van het onderzoek dat we hebben uitgevoerd in dit proefschrift, een overzicht van de actuele stand van zaken met betrekking tot het aanleren van technische en niet-technische vaardigheden tijdens chirurgische, 'hands-on' trainingssituaties van verschillende complexiteit. We beschrijven hierin ook hoe het aanleren van technische en niet-technische vaardigheden geïntegreerd kan worden in hedendaagse chirurgische opleidingen. We geven een theoretisch kader dat daarbij kan helpen. Dit hoofdstuk wordt afgesloten met onze visie op nuttig toekomstig onderzoek naar effectieve chirurgische coaching.

9

APPENDICES

DANKWOORD

Dan is het nu tijd voor het dankwoord. Het stuk waar ik al die jaren naar uitgekeken heb om te mogen schrijven. Iedereen die mij geholpen heeft, elk op zijn of haar manier, in meerdere of mindere mate, bewust of onbewust, mag ik hier een plekje geven. Te beginnen met mijn begeleidingsteam: Harry, Lia en Jan-Maarten.

Harry, dankjewel dat ik zo veel van jou heb mogen leren. Voor ik jou leerde kennen wist ik niet eens dat je ook buiten het voetbal kon promoveren. Jouw daadkracht, ideeën en vindingrijkheid heb ik altijd bewonderd. Onder jouw toeziend oog mocht ik mijn werk regelmatig presenteren in Amerika. Dit waren prachtige ervaringen. Op één van die congressen gaf je aan dat ik nu toch echt eens aan de vrouw moest. 'Laat hem maar praten' dacht ik bij mijzelf. Moet je nu eens kijken wie hier naast mij staat! Bedankt dat je mijn, maar zeker ook haar promotor wilde zijn.

Lia, wat ben jij een kei van een copromotor. Niet alleen ben je onwijs kundig en sta je garant voor goede input en ideeën, je bent ook een fantastisch mens. Te allen tijde kon ik bij jou terecht, of het nu over onderzoek ging, of over totaal iets anders. En dat is nog steeds zo. De laatste anderhalf jaar was je mijn leidinggevende binnen de Radboudumc Health Academy. Ik heb altijd ontzettend graag voor jou gewerkt. Binnenkort ga je het leidinggevende stokje overdragen. Ook al zal ik jou nog vaak tegenkomen, ik weet zeker dat ik je ga missen. Bedankt voor alles.

Jan-Maarten, wat was ik blij telkens als ik samen met jou achter een computer kon kruipen om naar mijn statistiek of schrijfsels te kijken. Je bent een bedreven onderzoeker die bol staat van de creativiteit en originele ideeën. Daar heb ik veel van geleerd. Ik waardeer je droge humor en nuchtere kijk op de zaak. Elke zaak. Ook wij zijn vaak samen op pad geweest in het buitenland. Dat zijn mooie herinneringen geworden. Bedankt dat je mijn copromotor wilde zijn.

De manuscriptcommissie: Thijs Merkx, Jacqueline de Graaf en Jos van der Hage. Hartelijk dank voor jullie tijd en energie waarmee jullie mijn proefschrift onder de loep hebben genomen. Ik kijk uit naar de verdediging.

Roland, wat heb ik stiekem veel aan jou als mentor gehad. Zo nu en dan kwam ik bij je langs om de voortgang van mijn onderzoek te bespreken. Je liet me mijn verhaal doen en luisterde aandachtig, ook tussen de regels door. Vervolgens hield je me vakkundig een spiegel voor. Waarom stelde ik bepaalde dingen nou toch zo uit? Nog steeds hoor ik jou die vraag af en toe stellen.

Edward en Marit, dank jullie wel dat ik steeds weer welkom was bij de Definitive Surgical and Anaesthetic Trauma Care cursus. Zonder al die data die we daar verzameld hebben was dit boekje er nooit gekomen. Alex, Wilma en Maikel, fijn dat jullie altijd zo vriendelijk en behulpzaam waren tijdens die drukke cursusdagen. Hartelijk dank ook aan al die instructeurs, cursisten, operatieassistenten, medeonderzoekers, studenten en vrienden die telkens weer meehielpen aan het verzamelen van die data op de manieren die wij aan de tekentafel hadden uitgedacht.

Alle medeonderzoekers van Heelkunde, dank voor de fijne werksfeer en het prettige arbeidsethos. Roger, Edwin en Marja, wat fijn dat ik af en toe binnen mocht lopen in het Heelkundelab. Die olifantentandpasta staat in mijn geheugen gegrift. Mats, dankjewel voor je nuttige bijdragen in de verwerking van de verzamelde stressdata.

Charles en Jos, helemaal in het begin van mijn onderzoek ging mijn harde schijf stuk. Jullie waren destijds het lichtpuntje aan het einde van die groene grot die mijn data wisten te redden. Bedankt. Geert, wat fijn dat ik ook af en toe met jou mocht statistieken. Dank voor je luchtige maar nuttige kijk op al die data. Collega's van de Radboudumc Health Academy, bedankt voor alle gezelligheid en mogelijkheden die jullie mij hebben geboden om stoom af te blazen.

Rik en Marieke, medepromovendi van onderwijskundige pluimage zijn dun gezaaid in het Radboudumc. Zeker die van jullie soort. Allerlei stukken hebben jullie bekeken en van feedback voorzien, jullie zaten boordevol tips over de organisatorische kant van een promotie en het bijbehorende boekje en jullie hebben mij ontzettend geholpen in de voorbereiding op mijn verdediging. Bedankt dat ik zo vaak bij jullie aan mocht kloppen.

Dan de helden van 3.47, dat toen nog jeugdige clubje onderzoekers in dat weggestopte hokje bij het oude kantoortje van Harry. Anders dan de omschrijving doet vermoeden was dit een ontzettend fijne werkplek. Dat vliegertje aan die ventilator, het real-life Mariospel met bureaustoelen en die late stapavonden zal ik nooit vergeten. Marjan, Chema, Dagmar en Willem, jullie zijn onwijs belangrijk geweest! Hetty, Dick, en Tjarda, fijn dat jullie af en toe de andere kant op wilden kijken.

Willemijn, al sinds de eerste statistiekopdrachten tijdens onze studie ben jij mijn trouwe squashmaatje. Ik weet zeker dat die Tripel Karmelietjes na het sporten hun bijdrage hebben geleverd, evenals de begeleidende gesprekken natuurlijk.

Dank aan alle voetballers en oud-voetballers van C.F.C. Brakkenstein. Al sprak ik zelden over mijn onderzoek, er zat vaak meer achter die ongecontroleerde schoten dan jullie konden vermoeden.

Malvertse Boys: Koen, Stijn, Dennis en Willy. Ook al hadden jullie dat misschien niet door, jullie zijn onmisbaar geweest. Wat begon als een bij elkaar geraapt groepje studenten dat elkaar gevonden had in hun zoektocht naar goedkope huisvesting, is uitgemond in een fantastische vriendengroep. Met jullie kon en kan ik mijn hoofd altijd leegmaken.

Tim, ik weet nog goed dat we elkaar op de fiets bij het Keizer Karelplein voor het eerst tegenkwamen. We waren beiden op weg naar de eerste introductiedag van onze nieuwe studie. 'Wat veel vrouwen hier!' Dat waren ongeveer de eerste woorden die we wisselden. Sindsdien hebben we als vrienden, studiegenoten, voetbalmaten en collega's al veel mooie herinneringen gemaakt. Bedankt dat jij mijn paranimf wilde zijn.

Stijn, wat een fijne vriend ben jij. Van een Groot Hart over de Neude tot aan de Malvert Mavericks met 40.000 man publiek. Maar bovenal sta jij altijd voor anderen klaar. Wij hebben al veel mooie momenten mogen beleven samen en daar zullen er ongetwijfeld nog veel aan worden toegevoegd. 'Ja, maar hoevéél?' Laten we beginnen met vandaag. Wat ongelooflijk tof dat jij mijn paranimf wilde zijn.

Luc, één les op de middelbare school hebben we naast elkaar mogen zitten. Van Soest zette ons direct apart. Sindsdien heeft onze vriendschap veel veranderingen doorstaan. Altijd heb jij de weg naar Nijmegen nog weten te vinden. Samen hebben we veel moois meegemaakt, in binnen- en buitenland. Laten we dat nog lang zo blijven doen.

Léon, toen we samen opgroeiden heb ik het nooit toegegeven. Tegenwoordig kan ik niet anders concluderen dan dat jij een hele fijne broer bent. Wij weten precies wat we aan elkaar hebben en zullen elkaar altijd weten te vinden.

Pap en mam, als kind gaven jullie mij mee dat ik alles mocht worden: van vrachtwagenchauffeur tot onderwijskundige, al wisten we van dat laatste beroep destijds het bestaan nog niet, zolang ik er maar mijn best voor deed. Dat heeft jullie altijd tot een hele fijne thuishaven gemaakt. En dan jullie onophoudelijke wil om te helpen. Met wat dan ook. Niets is jullie ooit te veel gevraagd. Bedankt voor alles!

Lieve Mariska, zonder enige twijfel ben jij het beste dat mijn promotie mij heeft opgeleverd. We begonnen als collega's, samen jagen op spuug van chirurgen om onderzoek mee te doen. Had jij ooit kunnen bedenken dat wij hier vandaag samen zouden staan? Wat heb ik ongelooflijk veel aan jou gehad. Helemaal toen mijn hoofd mij zo in de steek liet na die uit de hand gelopen trauma capitis (hersenschudding klinkt zo lullig). Jij was er altijd voor mij, van luisterend oor tot schop onder mijn kont. We hebben heel veel mooie toekomstplannen samen. Er is niets waar ik meer naar uitkijk dan deze met jou te gaan verwezenlijken. Ik hou van jou!

CURRICULUM VITAE

Sander (Alexander Paulus Beatrix) Alken was born on November 23, 1984, in Geleen, The Netherlands. After graduating from High School in 2003, he attended the Teacher Training for Primary Education at the Fontys University of Applied Sciences in Roermond. He obtained his Teacher Degree in 2007. That same year, he started to study Educational Science at the Radboud University Nijmegen. In 2011 he obtained a Bachelor and a Master Degree in Educational Science (both Bene Meritum). During his Master Internship he worked as an intern at the Academic Educational Institute of the Radboud University Medical Center in Nijmegen. Here he developed an interest in the coaching of surgical trainees by surgical teachers. In 2012 he started to work as a researcher at the Department of Surgery of the Radboud University Medical Center under the supervision of prof. dr. Harry van Goor (Professor of Surgical Edcuation). In 2014 he became a PhD candidate.

Sander lives together with his girlfriend Mariska Weenk in Nijmegen. He is currently working as an Educationalist at the Radboudumc Health Academy at the Radboud University Medical Center.

LIST OF PUBLICATIONS

2013	Feedback activities of instructors during a trauma surgery course. Alken A, Tan E, Luursema JM, Fluit C & van Goor H	*The American Journal of Surgery, 2013, 206(4): 599-604.*
2015	Coaching during a trauma surgery team training: perceptions versus structured observations. Alken A, Tan E, Luursema JM, Fluit C & van Goor H	*The American Journal of Surgery, 2015, 209(1): 163-169.*
2015	When experts are oceans apart: comparing expert performance values for proficiency-based laparoscopic simulator training. Luursema JM, Rover MM, Alken A, Kengen B & van Goor H	*The Journal of Surgical Education, 2015, 72(3):536-541.*
2017	Stress measurements in surgeons and residents using a smart patch. Weenk M, Alken A, Engelen L, Bredie S, van de Belt T & van Goor H	*The American Journal of Surgery, 2017, S0002-9610(16)31067-4.*
2017	Van camera's op de OK valt veel te leren. Alken A, Kriens M, van Goor H & Scheffer GJ	*Medisch Contact, 2017, www.medischcontact.nl*
2018	Integrating technical and non-technical skills coaching in an acute trauma surgery team training: Is it too much? Alken A, Luursema JM, Weenk M, Yauw S, Fluit C & van Goor H	*The American Journal of Surgery, 2018, 216(2):369-374.*
2018	Integrating technical and non-technical skills in hands-on surgical training. Alken A, Fluit C, Luursema JM & van Goor H	*Medical and Surgical Education: Past, Present and Future, 2018, DOI:10.5772/intechopen.73044*
2019	Measuring stress and coaching behaviors during a highly realistic trauma surgery team training. Alken A, Fluit C, Weenk M, Koeneman M, Luursema JM & van Goor H	*Submitted.*
2019	Evidence-based effective teaching behaviors for complex psychomotor skills training. Alken A, Luursema JM, Thornblade L, Chen X, Hull L, Horvath K, van Goor H & Fluit C	*Submitted.*

A

Institute for Health Sciences
Radboudumc

PHD PORTFOLIO

Name PhD candidate:	A.P.B. Alken	**PhD period:**	01-09-2014 – 02-11-2017
Department :	Surgery	**Promotor:**	Prof. Dr. H. van Goor
Graduate School:	Radboud Institute for Health Sciences	**Co-promotors:**	Dr. C.R.M.G. Fluit Dr. J.M. Luursema

	Year(s)	ECTS
TRAINING ACTIVITIES		
a) *Courses & Workshops*		
- Introduction Course for PhD students	2014	1.75
- Academic Writing Course	2014	3.00
- Scientific Integrity Course	2014	0.40
- English as a Second Language Course, University of California, Irvine, USA	2014	20.0
		25.15
b) *Symposia & Congresses**		
- Association of Surgeons of Great Britain and Ireland Congress, Harrogate, UK, poster presentation	2014	1.00
- Dutch Society of Obstetrics and Gynaecology Congress, Papendal, The Netherlands, podium presentation	2014	0.50
- Association for Surgical Education - Surgical Education Week Congress, Seattle, USA, attendee	2015	1.25
- International Association of Medical Science Educators Congress, San Diego, USA, poster presentation	2015	1.25
- Netherlands Association for Medical Education Congress, Rotterdam, The Netherlands, poster presentation	2015	0.75
- Association for Surgical Education - Surgical Education Week Congress, Boston, USA, podium presentation	2016	1.75
- Netherlands Association for Medical Education Congress, Egmond aan zee, The Netherlands, attendee	2017	0.50
- Association for Surgical Education - Surgical Education Week Congress, San Diego, USA, attendee	2017	1.25
- Academic Surgical Congress, Las Vegas, USA, Quick shot presentation	2017	1.25
		9.5
c) *Other*		
- Monthly Research Meetings, Department for Research in Learning and Education, Radboudumc Health Academy	2014-17	3.00
TEACHING ACTIVITIES		
d) *Lecturing / teaching*		
- Teach the Teacher Training Course for Medical Interns, Master Phase of Medicine, CK08; every month a two-day course	2014-17	57.60
- Basic Laparoscopic Skills Course for Medical Interns, Master Phase of Medicine, CK03; two hours per week	2014-17	62.40
		120.0
e) *Supervision of internships / other*		
- Meet your PhD, Biomedical Sciences Bachelor Phase	2017	0.40
TOTAL		**158,05**

*Indicate oral or poster presentation

boud University Radboudumc university medical center

A

Institute for Health Sciences
Radboudumc