

Incorporation of proficiency criteria for basic laparoscopic skills training: how does it work?

E. G. G. Verdaasdonk · J. Dankelman ·
J. F. Lange · L. P. S. Stassen

Received: 14 May 2007 / Accepted: 22 December 2007 / Published online: 4 April 2008
© The Author(s) 2008

Abstract

Background It is desirable that surgical trainees are proficient in basic laparoscopic motor skills (eye–hand coordination). The present study evaluated the use of predefined proficiency criteria on a basic virtual reality (VR) simulator in preparation for a laparoscopic course on animal models.

Methods Twenty-eight surgical trainees who enrolled for a basic laparoscopic course were trained on a basic (VR) simulator until their performance met predefined criteria. Two different criteria were defined, based on the performance of experienced laparoscopic surgeons on the simulator. In the first group ($n = 10$), the criteria were set at the 75th percentile of the laparoscopic surgeons' performance on the simulator and in the second group, at the 50th percentile ($n = 18$). Training time and number of attempts needed until the performance criteria were met were measured.

Results In the first group, training time needed to pass the test ranged from 29 to 77 min (median: 63 min) with a

range of 43–90 attempts (median 61 attempts). In the second group, training time ranged from 38 to 180 min (median 80 min) with a range of 55–233 attempts (median 95 attempts). Experience with assisting or performing laparoscopic procedures varied widely and was not correlated with the training time and number of attempts needed to pass the criteria.

Conclusions The performance criteria for training laparoscopic motor skills on a (VR) simulator resulted in wide variation between surgical trainees in time and number of attempts needed to pass the criteria. This demands training courses with a flexible time span tailored to the individual level of the trainee.

Keywords Training · Education · Surgery · Laparoscopic · Virtual reality · Simulation · Proficiency · Criterion-based

Laparoscopic training courses have become important education tools in the surgical curriculum. Their ultimate aim is to reduce the learning curve of surgeons on actual patients. Most courses take two or more days and consist of lectures, laparoscopy videos and, most importantly, motor skills training. Training of laparoscopic motor skills is a central part of these courses.

VR trainers have become attractive and valuable tools to train surgeons in a non-patient non-animal environment [1, 2] and proven effective in learning basic skills that can be transferred to real procedures [3, 4]. In particular, the automated assessment, feedback and unlimited use of standardized tasks seem to offer advantages. However, due to new working-hour directives of 80 h per week in the USA and 48 h in Europe by 2011, training courses must not only be effective but also efficient. The learning effect should be maximal while using the least amount of time.

E. G. G. Verdaasdonk (✉)
Faculty of Mechanical, Marine and Materials Engineering,
Department of BioMechanical Engineering, Delft University of
Technology, Mekelweg 2, Delft 2628 CD, The Netherlands
e-mail: e.g.verdaasdonk@tudelft.nl

E. G. G. Verdaasdonk · J. Dankelman · L. P. S. Stassen
Man-Machine Systems Group, Department of BioMechanical
Engineering, Delft University of Technology, Delft, The
Netherlands

E. G. G. Verdaasdonk · L. P. S. Stassen
Department of Surgery, Reinier de Graaf Group, Delft, The
Netherlands

J. F. Lange
Department of Surgery, Erasmus Medical Center Rotterdam,
Rotterdam, The Netherlands

Most training courses employ a fixed duration for training while learning may be highly dependent on individual characteristics such as innate ability, previous experience, and motivation. In an optimal training course the available training time should be tailored to the individual level of the trainee. Therefore, proficiency of a given skill should depend on passing a clear criterion rather than an arbitrary amount of time or repetitions. Surgical society has only recently entered the realm of criterion-based training. Standards on which to define training endpoints are receiving attention [5] and some studies indicated the benefits of using preset proficiency criteria [6, 7]. Implementation of proficiency criteria may allow trainees to reliably achieve maximal benefit while minimizing unnecessary training [8]. However, the practical consequences of using such endpoints on course design and time schedule are unclear.

The aim of the present study was to assess the consequences of a criterion-based training program to train basic laparoscopic surgery skills using a VR simulator.

Two different proficiency criteria levels, based on the performance of experienced laparoscopic surgeons, were applied to train eye–hand coordination: easy versus difficult. The feasibility, usefulness, and challenge of these levels were

evaluated. The potential consequences and difficulties of defining performance criteria for VR simulators for the purpose of recruitment, selection, and licensing are discussed.

Methods

In 2006, all the surgical trainees who enrolled for the basic laparoscopic skills course in Rotterdam (Skills Centre, Erasmus Medical Centre Rotterdam, The Netherlands) were required to achieve an expert-based proficiency criterion on a VR simulator prior to embarking on animal models. Trainees were allowed to train as long as they needed in their spare time until the proficiency criterion was reached. Training took place within 2 weeks prior to the training on the animal model.

The SIMENDO[®] endoscopic simulator (Delltech, Delft, the Netherlands) was used. This basic simulator aims specifically at eye–hand coordination training applicable to laparoscopic surgery and employs abstract tasks without force feedback [9]. The cost of the simulator hardware and software was about 9,000 EUR, excluding the desktop computer. Six tasks were selected from the simulator software SimSoft 1.0 (Delltech, Delft, the Netherlands)

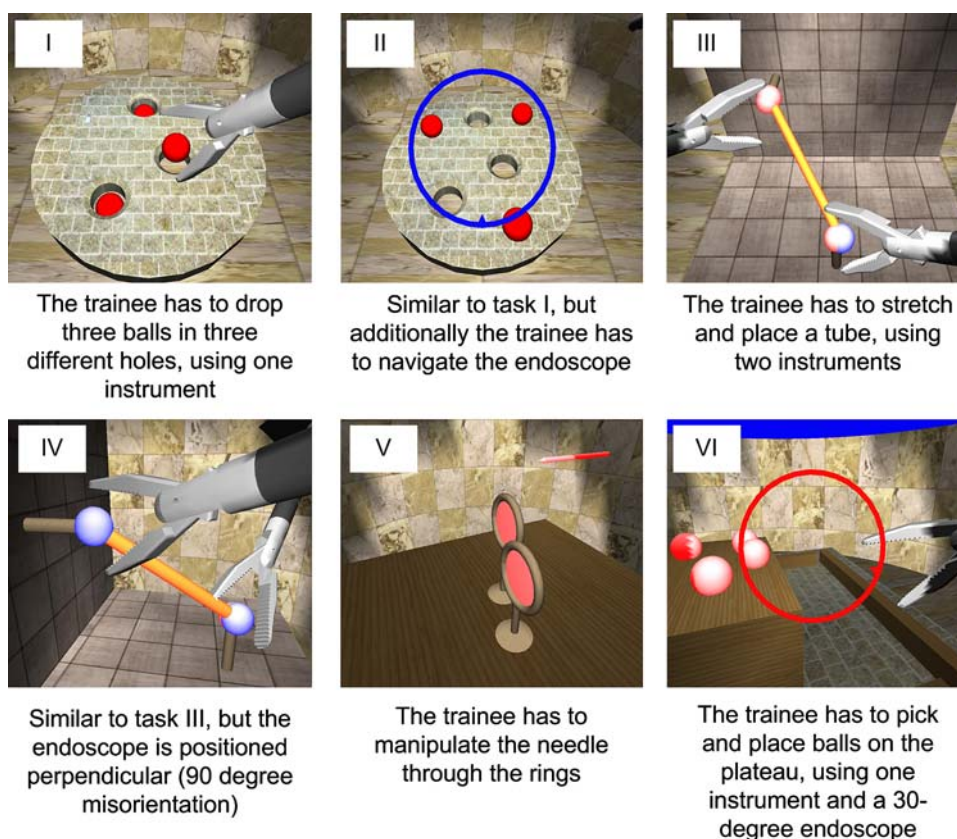
Table 1 Tasks description and results of the two groups with different proficiency criteria levels

Task name	Parameter	Group 1 (75th percentile ^a)	Group 2 (50th percentile ^a)
I. Drop the balls (single instrument)	Time (s)	21	17
	Collision (number)	1	0
	R path length (a.u.)	24	21
II. Drop the balls & endoscope	Time (s)	23	18
	Collision (number)	1	0
	R path length (a.u.)	22	19.5
	Endoscope movement (a.u.)	5.5	2.5
	Camera aim (%)	75	75
III. Stretch (no misorientation)	Time (s)	15	12
	Collision (number)	1	0
	R path length (a.u.)	11	9
	L path length (a.u.)	13	11
IV. Stretch (90° misorientation)	Time (s)	25	17
	Collision (number)	1	0
	R path length (a.u.)	20	14
	L path length (a.u.)	18	13
V. Ring & needle	Time (s)	33	28
	Collision (number)	3	2
	R path length (a.u.)	28	24
	L path length (a.u.)	22	19.5
VI. 30° endoscope & pick and place	Time (sec)	41	33
	Collision (number)	1	0
	R path length (a.u.)	65	61
	Endoscope movement (a.u.)	30	23
	Camera aim (%)	75	75

a.u., arbitrary units

^a Proficiency level in each group

Fig. 1 The six exercises and descriptions from the VR simulator, corresponding with the names in Table 1



(Table 1, Fig. 1). Measurement parameters were: time to complete the task, collisions with the nontarget environment, instrument path length, and aiming the endoscope. Correct aim was defined as the percentage of time that the endoscope was centered on the tip of the laparoscopic instrument. The proficiency levels were derived from previous work in which the same six tasks had been executed by 15 experienced laparoscopic surgeons (more than 50 laparoscopic procedures performed) [10]. Our proficiency criteria were based on the fifth repetition of the tasks.

Two different proficiency criteria levels were defined. In the first group of ten trainees (group 1), the predefined levels were set at the 75th percentile (easy) of the experts' proficiency on the VR simulator. In the second group of 18 trainees (group 2), the 50th percentile or median (difficult) was employed. Table 1 displays the values required for each task. To pass the test the trainees had to continue training until this level of proficiency was reached in three consecutive repetitions. The total training time (including breaks) and the number of attempts they needed to pass the criteria was measured.

The trainees signed an informed consent to use the data for scientific research and filled in a questionnaire about their previous experience with laparoscopic surgery and laboratory training. After successfully completing the training on the VR simulator, all trainees answered seven

questions about the usefulness, feasibility, and challenge of the training at their particular preset proficiency level. Furthermore, they scored the degree of challenge on a scale from 1 (none) to 10 (enormous).

Results

Table 2 shows the median number of attempts (and ranges) needed to pass the preset level per task. The total number of repetitions needed to pass the proficiency level also varied widely between the individual trainees within each groups. In the first group at the 75th percentile level of the experts, the training time needed to pass the test ranged from 29.4 to 77.1 min (median: 63.9 min) with a range of 43–90 attempts (median 61 attempts). In the second group (50th percentile level), training time ranged from 37.8 to 179.9 min (median 80 min) with a range of 55–233 attempts (median 95 attempts). All the trainees accomplished the predefined criteria.

In group 1 (75th percentile), the fastest 25% of the trainees needed fewer than ten repetitions to pass the criteria for the 30° endoscope navigation and delicate needle-handling tasks. The slowest 25% needed more than 15 repetitions. In group 2 (50th percentile), the fastest 25% of the trainees needed fewer than 14 repetitions to pass the

Table 2 Number of attempts needed per task to reach the proficiency level for each group

Task description	Group 1 (75th percentile ^a) Median attempts (range)	Group 2 (50th percentile ^a) Median attempts (range)
I. Drop the balls (single instrument)	9 (4–18)	17 (5–58)
II. Drop the balls & endoscope	7 (4–22)	13 (3–38)
III. Stretch (normal camera view)	9 (5–13)	11 (4–50)
IV. Stretch (difficult camera view)	6 (3–10)	13 (5–40)
V. Ring & needle	13 (5–38)	20 (6–88)
VI. 30° endoscope & pick and place	13 (7–22)	23 (10–46)
Total attempts	61 (34–90)	95 (55–233)

Group 1 ($n = 10$) and group 2 ($n = 18$)

^a Proficiency level in each group

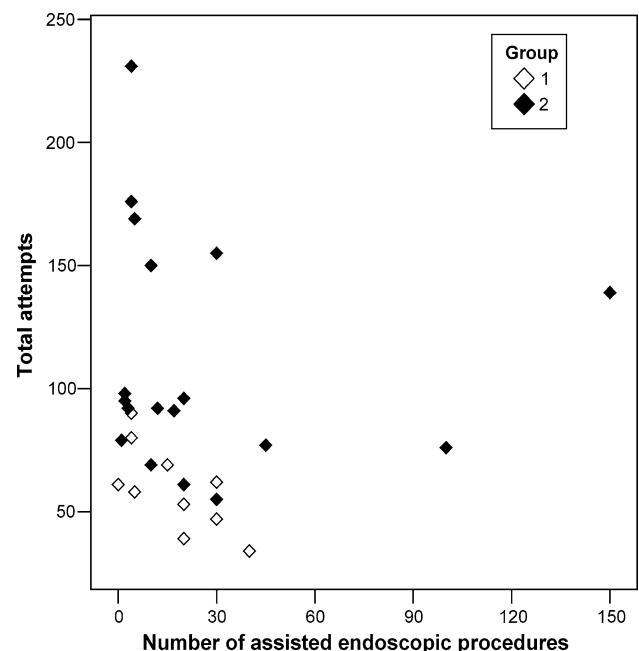
Table 3 Characteristics of the surgical trainees in groups 1 and 2

	Group 1 ($n = 10$)	Group 2 ($n = 18$)
Male: female	10:0	12:6
Median age (range)	28.5 (28–30)	29.5 (26–32)
First year of surgical training (n)	5	7
Second year of training (n)	5	11
Previous experience		
Lab training in a box or VR (n)	3	4
Median number of endoscopic procedures		
Assisted	17.5 (5–40)	11 (1–150)
Partially performed	0.5 (0–8)	0 (0–10)
Completely performed	0 (0–6)	0.5 (0–15)

two tasks compared to more than 25 repetitions in the slowest 25% of the trainees in this group.

Experience with assisting with or performing endoscopic procedures (under supervision) and training in laboratories varied widely between the trainees. Table 3 shows the characteristics and experience of the trainees. Figure 2 shows the relationship between the number of endoscopic procedures that the trainees assisted and the number of attempts needed to pass the criteria. No statistically significant correlation was found between these two parameters in either of the groups. There was also no significant correlation between the number of endoscopic procedures performed under supervision by the trainees and the number of attempts (Fig. 3). However, none of the trainees who had performed more than two endoscopic procedures needed more than 100 attempts.

The results of the questionnaire, filled in after the training, are shown in Table 4. In the 75th percentile group (group 1), three out of the ten participants stated that the criteria were too easy (30%), whereas none of the participants in the 50th percentile group (group 2) found their level was too easy. The two groups rated the challenge of the training program with a score of 8.

**Fig. 2** Number of endoscopic procedures assisted and number of attempts needed to achieve the predefined proficiency criteria for group 1 (75th percentile) and group 2 (50th percentile)

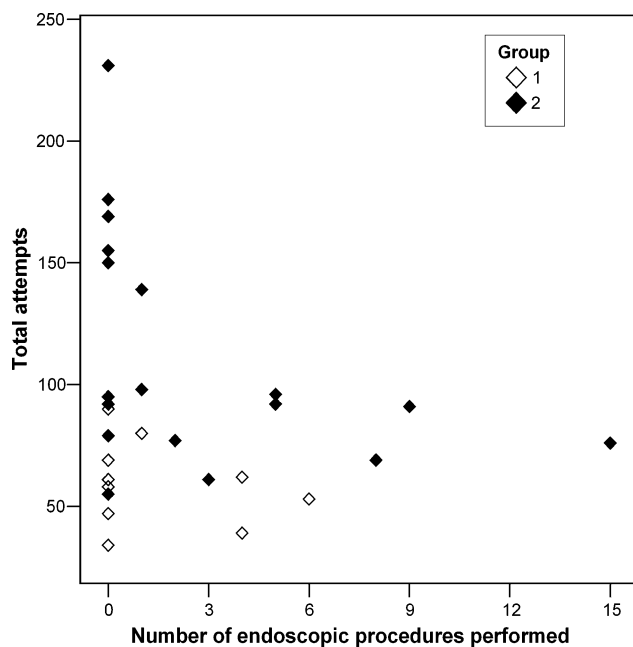


Fig. 3 Number of endoscopic procedures performed and number of attempts needed to achieve the predefined proficiency criteria for group 1 (75th percentile) and group 2 (50th percentile)

Discussion

The criterion-based motor skills training program on our VR simulator was intended to prepare surgical trainees efficiently for a laparoscopic course on a porcine model. The total number of repetitions needed to pass the proficiency level varied widely between the individual trainees within each group. Some trainees needed up to four times more time to pass the test than others.

All the trainees found the training useful and were able to achieve the predefined criteria, which meant that the set-up had good feasibility. Setting the proficiency criteria at a more difficult level (50th percentile of the median expert score) appeared to be more appropriate than the easier setting (75th percentile).

In the literature, there is growing evidence that motor skills training with inanimate VR simulators is valid and improves the performance of actual procedures [11]. However, it is not yet clear how these training models programs should be standardized and incorporated into the a surgical curriculum. Physical simulators or box trainers have shown to train laparoscopic motor skills effectively as well [12] and competence levels based on performance of experienced laparoscopic surgeons seem suitably challenging for novices [8]. Nevertheless, task time was the only parameter used. Another study successfully trained novices laparoscopic suturing by using an expert performance level, based on time and error assessment [6]. However, the advantage of VR simulators is that

performance is measured, stored, and displayed automatically. Furthermore, most VR simulators employ several parameters to assess performance, such as task time, collisions with the VR environment, instrument path length, and numbers of specifically defined errors. However, there is an ongoing debate about the pass or fail standards of these parameters. In general, the concept of criterion-based training aims to introduce standards that provide surgical educators with strategies to design a transparent and validated training program. Evaluation of the experimental set-up provides insight into the feasibility of the tasks, the performance criteria, and practical issues such as the duration of training.

Several remarks should be made about this new concept of criterion-based skills training. To validate a criterion-based training program on a simulator and incorporated it into a structured surgical curriculum the following requirements should be met: (1) the goal of the simulator training program has to be defined and validated in terms of which skills are actually learned, (2) the performance criteria have to be determined based on experienced surgeon performance on the simulation and evaluated to offer trainees straightforward and challenging exercises, (3) inexperienced trainees should be able to meet the criteria, and the consequences of failing to achieve the required criteria level should be made clear beforehand. The student is allowed to progress to more advanced training setting when the criterion is achieved. Students who do not meet the criteria should receive more training, feedback, and retest opportunities.

The VR simulator and tasks used in this study were aimed specifically at teaching trainees the basic motor skills needed for laparoscopic surgery (e.g., hand–eye coordination). Previous studies showed that the simulator had content and construct validity [9, 10]. The VR simulator is a valid model at the beginning of the learning curve in laparoscopic surgery.

Determining performance criteria for VR simulator training is more difficult than it seems. In the literature there is no consensus on how to define these criteria. Some authors advised that a certain number of repetitions should be performed [13], whereas others recommended the use of preset criteria rather than a predetermined training duration or an arbitrary number of repetitions [5, 7, 14]. When preset criteria and corresponding scores are chosen, the exact scores depend on the researcher and type of simulator. In a study on the minimally invasive surgical trainer virtual reality (MIST-VR) simulator [15], the authors remarked that, if performance criteria are based on the average scores of experienced laparoscopic surgeons, the level might be too easy. Instead, they recommended using the score of an experienced laparoscopic surgeon who also had extensive experience on the simulator. Aggarwal et al.

Table 4 Results of the questionnaire filled out by the participants after the simulator training

Questions	Group 1 (<i>n</i> = 10) (75th percentile ^a)	Group 2 (<i>n</i> = 18) (50th percentile ^a)
1. What is your opinion about the competence criteria?		
Too difficult	0%	5.50%
Good	70%	94.50%
Too easy	30%	0%
2. What is your opinion about training on this simulator in general?		
Not useful	0%	0%
Useful	100%	100%
No opinion	0%	0%
3. What is your opinion about the tasks?		
Poor	0%	0%
Good	90%	100%
No opinion	10%	0%
4. What is your opinion about the feedback parameters?		
Poor	0%	0%
Good	90%	94%
No opinion	10%	6%
5. Is the training challenging? Give a score between 1 and 10 (10 = greatest challenge)	median 8 (range 6–9)	median 8 (range 5–10)
6. Does the simulator training improve the eye–hand coordination of inexperienced trainees?		
No	0%	0%
Yes	100%	100%
No opinion	0%	0%
7. Is the simulator capable of objectively measuring skills that are important to laparoscopic surgery?		
No	0%	0%
Yes	100%	78%
No opinion	0%	22%

^a Proficiency level in each group

used the median score achieved by ten experienced surgeons in two consecutive repetitions in a study on the LapSim simulator [14].

In the first group the 75th percentile of the expert performance was chosen because the 50th percentile (median) performance scores were expected to be too hard to acquire by the average trainee. However, 30% of the residents considered the 75th percentile too easy. In the next group, only one participant thought the 50th percentile was too difficult. Therefore, the 50th percentile is considered to be more useful.

Introducing criterion-based training motor skills training raises questions about the validity of the criteria for the recruitment and selection purposes of trainees, or (re)licensing of surgeons. Setting the criteria at the median of experts' performance means that, by definition, 50% of the experts do not achieve this level either. Therefore, this level cannot be used for high-stake examination of surgeons, e.g. (re)licensing, but it may be justified for recruitment and selection purposes of inexperienced trainees. Obviously, passing a motor skills test on its own does not guarantee that the individual is competent in all the

required domains. Good motor skills are only one of the necessary requirements to become a competent laparoscopic surgeon [16, 17]. On the other hand, if a trainee is unable to pass a validated simulation test or demonstrate improvement during training, then a surgical career is questionable. Our results revealed that the current criteria form an efficient means to shape the hand–eye coordination of those who need it and enhance the process of skills acquisition, an essential prerequisite of high-standard surgery. Although the SIMENDO[®] forms a valid model to train subjects with little or no experience with laparoscopic surgery, it seems less suitable for general performance assessment of experienced laparoscopic surgeons for licensing purposes. These high-stake examinations require more complex simulation programs, or combinations of a battery of different test modules, for example, programs that make objective evaluations of decisional behavior, proper reactions on adverse events, anatomical knowledge, etc. and thus test competence on a broader scale.

An important practical issue is that the consequences of not passing the test on the simulator must be made clear beforehand. In this study, all the trainees achieved the

predefined criteria. Part of the exercise was to train until they met the performance criteria, even if this took a great many training sessions. For future training programs, it is expected that most surgical trainees will be able to pass the set criteria. This assumption is based on the observation that surgical trainees are highly motivated to learn the required skills and invest the necessary time. In addition, there seems to be natural selection within the surgical population itself. However, this assumption must be considered cautiously, because this might not apply for other simulators and the assessment of subjects with different motivation, interests, and backgrounds. Schijven et al. [18] found that in a clip-and-cut task on an advanced VR simulator, 20% of the 30 participants could not improve their performance score sufficiently to obtain proficiency in 30 repetitions. However, not all participants in the study were surgical trainees. The participants comprised of a mixed group of final-year medical students, internal medicine trainees, trainees in the department of anesthesia, and surgical trainees. This might suggest that the selected population of surgical residents, as in our study, is more likely to pass the set criteria. Furthermore, a study by Brunner et al. [5] that used basic exercises on the MIST-VR indicated that a lengthy learning curve existed for novices, possibly beyond 30 repetitions. In their opinion, performance plateaus may not reliably determine training endpoints [5].

In conclusion, criterion-based training of motor skills on a VR simulator is an efficient, feasible, and useful method to prepare surgical trainees for more complex procedures, for example, on animal models. Median expert performance scores seemed appropriate as proficiency criteria. The use of the criteria resulted in wide variation between surgical trainees in time and number of attempts needed to pass the criteria. Therefore, it is particularly suitable for the selection of trainees who need more basic motor skills training and providing them with enough time to acquire these skills. Consequential, training programs could become more effective if tailored to the individual's level. Such flexible courses are currently not common in surgical training.

Acknowledgement The research was financially supported by a grant provided by the health insurance company DSW (Schiedam, The Netherlands).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Jakimowicz JJ, Cuschieri A (2005) Time for evidence-based minimal access surgery training—simulate or sink. *Surg Endosc* 19(12):1521–1522
- Champion HR, Gallagher AG (2003) Surgical simulation—a 'good idea whose time has come'. *Br J Surg* 90(7):767–768
- Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P (2004) Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 91(2):146–150
- Hyltander A, Liljegren E, Rhodin PH, Lonroth H (2002) The transfer of basic skills learned in a laparoscopic simulator to the operating room. *Surg Endosc* 16(9):1324–1328
- Brunner WC, Korndorffer JR Jr, Sierra R, Massarweh NN, Dunne JB, Yau CL, Scott DJ (2004) Laparoscopic virtual reality training: are 30 repetitions enough? *J Surg Res* 122(2):150–156
- Korndorffer JR Jr, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ (2005) Simulator training for laparoscopic suturing using performance goals translates to the operating room. *J Am Coll Surg* 201(1):23–29
- Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A (2006) A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg* 191(1):128–133
- Korndorffer JR Jr, Scott DJ, Sierra R, Brunner WC, Dunne JB, Slakey DP, Townsend MC, Hewitt RL (2005) Developing and testing competency levels for laparoscopic skills training. *Arch Surg* 140(1):80–84
- Verdaasdonk EG, Stassen LP, Monteny LJ, Dankelman J (2006) Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO. *Surg Endosc* 20(3):511–518
- Verdaasdonk EG, Stassen LP, Schijven MP, Dankelman J (2007) Construct validity and assessment of the learning curve for the SIMENDO endoscopic simulator. *Surg Endosc* 21(8):1406–1412
- Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB, Jakimowicz JJ (2005) Consensus guidelines for validation of virtual reality surgical simulators. *Surg Endosc* 19(12):1523–1532
- Scott DJ, Bergen PC, Rege RV, Laycock R, Tesfay ST, Valentine RJ, Euhus DM, Jeyarajah DR, Thompson WM, Jones DB (2000) Laparoscopic training on bench models: better and more cost effective than operating room experience? *J Am Coll Surg* 191(3):272–283
- Hackethal A, Immenroth M, Burger T (2006) Evaluation of target scores and benchmarks for the traversal task scenario of the Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR) laparoscopy simulator. *Surg Endosc* 20(4):645–650
- Aggarwal R, Grantcharov TP, Eriksen JR, Blirup D, Kristiansen VB, Funch-Jensen P, Darzi A (2006) An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg* 244(2):310–314
- Brunner WC, Korndorffer JR Jr, Sierra R, Dunne JB, Yau CL, Corsetti RL, Slakey DP, Townsend MC, Scott DJ (2005) Determining standards for laparoscopic proficiency using virtual reality. *Am Surg* 71(1):29–35
- Baldwin PJ, Paisley AM, Brown SP (1999) Consultant surgeons' opinion of the skills required of basic surgical trainees. *Br J Surg* 86(8):1078–1082
- Cuschieri A, Francis N, Crosby J, Hanna GB (2001) What do master surgeons think of surgical competence and revalidation? *Am J Surg* 182(2):110–116
- Schijven MP, Jakimowicz JJ, Carter FJ (2004) How to select aspirant laparoscopic surgical trainees: establishing concurrent validity comparing Xitact LS500 index performance scores with standardized psychomotor aptitude test battery scores. *J Surg Res* 121(1):112–119