

Criterion-based laparoscopic training reduces total training time

Willem M. Brinkman · Sonja N. Buzink ·
Leonidas Alevizos · Ignace H. J. T. de Hingh ·
Jack J. Jakimowicz

Received: 27 June 2011 / Accepted: 10 October 2011 / Published online: 1 November 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract

Introduction The benefits of criterion-based laparoscopic training over time-oriented training are unclear. The purpose of this study is to compare these types of training based on training outcome and time efficiency.

Methods During four training sessions within 1 week (one session per day) 34 medical interns (no laparoscopic experience) practiced on two basic tasks on the Symbionix LAP Mentor virtual-reality (VR) simulator: ‘clipping and grasping’ and ‘cutting’. Group C (criterion-based) ($N = 17$) trained to reach predefined criteria and stopped training in each session when these criteria were met, with a maximum training time of 1 h. Group T (time-based) ($N = 17$) trained for a fixed time of 1 h each session. Retention of skills was assessed 1 week after training. In addition, transferability of skills was established using the Haptica ProMIS augmented-reality simulator.

Results Both groups improved their performance significantly over the course of the training sessions (Wilcoxon signed ranks, $P < 0.05$). Both groups showed skill

transferability and skill retention. When comparing the performance parameters of group C and group T, their performances in the first, the last and the retention training sessions did not differ significantly (Mann–Whitney U test, $P > 0.05$). The average number of repetitions needed to meet the criteria also did not differ between the groups. Overall, group C spent less time training on the simulator than did group T (74:48 and 120:10 min, respectively; $P < 0.001$). Group C performed significantly fewer repetitions of each task, overall and in session 2, 3 and 4.

Conclusions Criterion-based training of basic laparoscopic skills can reduce the overall training time with no impact on training outcome, transferability or retention of skills. Criterion-based should be the training of choice in laparoscopic skills curricula.

Keywords Laparoscopic training · Criterion-based · Proficiency-based · LAP Mentor · Training time · Overtraining

Presented at the EAES 2011 Congress May 16–May 18, 2011 Torino, Italy.

W. M. Brinkman (✉) · I. H. J. T. de Hingh · J. J. Jakimowicz
Department of Surgery, Catharina Hospital Eindhoven,
Michelangelolaan 2, 5623 EJ Eindhoven, The Netherlands
e-mail: wnbrinkman@gmail.com

S. N. Buzink · J. J. Jakimowicz
Faculty of Industrial Design Engineering, Delft University
of Technology, Landbergstraat 15, 2628 CE Delft,
The Netherlands

L. Alevizos
Department of Surgery, Hippokraton Hospital,
114 Queen’s Sofia Avenue, Athens, Greece

Virtual-reality (VR) simulators are widely implemented in laparoscopic surgical training programmes to train psychomotor skills associated with this kind of surgery [1]. By training basic skills on a virtual-reality simulator, the first part of the learning curve of laparoscopic surgery is moved out of the operating room into the skills lab. Training basic laparoscopic skills in a skills-lab setting is proven to improve performance in the operating room [2, 3].

Different simulators have been produced and validated for training of basic laparoscopic skills [4, 5]. However, the optimal implementation of simulators in training programs remains a topic of discussion and investigation.

Due to the implementation of European legislation that reduced trainee working hours and the increased workload

due to rising use of healthcare facilities, training time needs to be used as efficiently as possible [6]. Therefore, it is important to know what is the most beneficial skills-lab training time, training schedule and training program. Research has shown, for instance, that an optimal training program is distributed over more days [7, 8]. Recent literature also suggests that the optimal endpoint for simulator training is the attainment of a predefined level (criterion-based training), rather than the completion of an arbitrary number of procedures, task repetitions or hours using the simulator (time-based training) [9, 10]. Also criterion-based training is supposed to boost resident motivation [11].

While the performance and motivational benefits of criterion-based VR simulator training have been demonstrated in previous studies, the training time benefits associated with such training are unclear [9–11].

The purpose of this study is to compare criterion-based training with time-based training to investigate whether criterion-based training is better than time-based training with respect to training outcome, transferability of skills, skills retention and training time.

Methods and materials

Protocol

In this study, 34 medical interns completed a simulator training program of four training sessions within 1 week (one session per day). In the introduction to the study it was explained to the participants that the researchers were not affiliated with the manufacturer of the simulator and that all data would be analysed anonymously. Informed consent was given by all participants ($N = 34$), after which they commenced the study by filling out a questionnaire about demographics and prior laparoscopic or laparoscopic simulation experience (Fig. 1). Subsequently, the participants watched a demonstration video about laparoscopic simulation and usage of the tools. They all started training on the ProMIS I or III augmented-reality (AR) simulator (Haptica, Dublin, Ireland) to determine a baseline performance level. The simulator displayed a demonstration video previous to the task, and step-by-step verbal explanation was given by the simulator during the training. All participants performed twice a translocation task and twice a sharp dissection task. The first exercise was to become familiar with the simulator; the second repetition was used to determine the baseline performance level. Thereafter, all participants received the same introduction on the LAP Mentor II (Symbionix Corp., Cleveland, USA) simulator by three informative posters. The participants performed the clipping and grasping task (task 5) and the cutting task

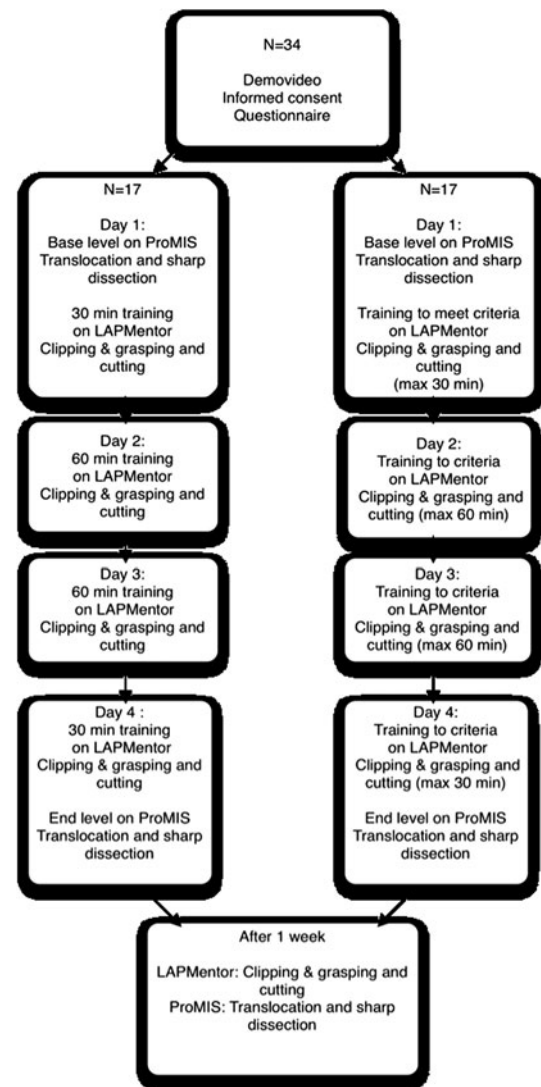


Fig. 1 The study protocol

(task 7) on the LAP Mentor. The second repetition on day 1 on the LAP Mentor was used to determine a baseline performance level, and the last repetition of each task on day 4 was used for the post-test. After training on day 4, a post-training performance level was established by performing twice the same translocation and sharp dissection task as on day 1 on the ProMIS (Fig. 1). The level of retention was established 1 week after training by performance of the two tasks on the LAP Mentor and the ProMIS. The participants, 34 medical interns in total (21 from Eindhoven, The Netherlands, and 13 from Athens, Greece), were randomly allotted to one of two groups. In the first group (group T, $N = 17$) the training was time based. Participants in group T performed the clipping and grasping task and the cutting task on the LAP Mentor for a fixed time period (Fig. 1). They completed four training sessions within 7 days on the LAP Mentor (180 min in total).

A 4-day training program with 1-h sessions (maximum) was chosen to ensure that the participants would overcome the initial learning curve [12] and assure that some over-training would take place. We divided training over multiple days to improve training performance [7] and to prevent exceeding a maximum of one training hour a day, the estimated maximum time besides an intern's mandatory clinical attendance.

The second group (group C, $N = 17$) trained on the LAP Mentor until their performances matched specific predefined performance criteria (Table 1). The criteria used in this study were derived from the performances of experienced surgeons [13]. When the participants in group C achieved the criteria on each task twice, they could stop training on that task for that day. The consecutive training day, they trained again until they achieved the criteria.

Equipment

The LAP Mentor is a VR-based laparoscopic training system. The software of the LAP Mentor II offers a variety of basic and procedural tasks in a VR environment to train different laparoscopy skills. After the performance of each task, the software provides numerical scores. In this study

two basic tasks were used: 'clipping and grasping' and 'cutting'.

The ProMIS augmented-reality (AR) simulator was used in this study to assess the transferability of the skills learned on the LAP mentor. The ProMIS AR simulator consists of a torso-shaped mannequin with a neoprene cover, containing an instrument tracking system. Different trays may be placed in the mannequin for each task, such as for the 'translocation' and the 'dissection' tasks we used in this study. The tasks are performed with AutoSuture disposable 5-mm Endo Clinch and Endo Shears (Covidien, Dublin, Ireland).

Statistics

The Dutch participants ($N = 21$) used a ProMIS I system, while the Greek participants ($N = 13$) trained on a ProMIS III system. Because of different data output settings, we analysed the two groups separately. All data were processed and analysed using the Statistical Package for the Social Sciences 18.0 (SPSS Inc., Chicago, USA). To analyse the differences in performances the Mann–Whitney U test (between the groups) and Wilcoxon signed-rank test (within the groups) were used. P value <0.05 was considered as statistically significant.

Table 1 Description of tasks and criteria

Simulator	Basic skill	Description of skill task	Criteria
LAP Mentor	Task 5, clipping and grasping	Grasp a leaking duct, stretch it until the red segment turns green and place a clip on the green segment	Time $\leq 1:41$ Clipped ducts = 9 Accuracy rate $\geq 90\%$ Eco of mov clipper $\geq 60.2\%$ Eco of mov grasper $\geq 54.1\%$
	Task 7, cutting	Retract the form and cut the fibres in a circle	Total time $\leq 1:30$ Accuracy rate = 100% Safe retraction $\geq 75\%$ Total path length right instr ≤ 251.3 cm Total path length left instr ≤ 83.3 cm
ProMIS	Instrument handling: object positioning	Pick up a number of objects, and place them in a specified target area	
	Dissection	Dissect a circle out of a stretched rubber glove	

Results

All participants ($N = 34$) in both of the groups improved their performances on the LAP Mentor tasks significantly over the course of the training sessions based on the parameters of time, economy of movement and path length. Figure 2 presents box plots of two parameters tested on the LAP Mentor: time and path length, for both tasks. Comparing the performance parameters of group C and group T, their performances in the first, the last and the retention training sessions did not differ significantly between the groups.

In both groups the skills acquired on the LAP Mentor transferred equally to their performances on the ProMIS; their performance on the ProMIS simulator improved between the pre-test and the post-test. Improvement was not significant for all tested parameters (Tables 2, 3). Participants in both groups showed skill retention in task performance on the LAP Mentor and ProMIS simulator. Performance on the retention test did not differ significantly from the post-test (Tables 2, 3).

Besides the performance metrics of both groups, we analysed the number of repetitions of tasks and the total time spent on the simulator. Group C performed significantly fewer repetitions of each task, overall and in session 2, 3 and 4 (Table 4). Altogether, group C spent significantly less time training on the simulator than group T (74:48 and 120:10 min,

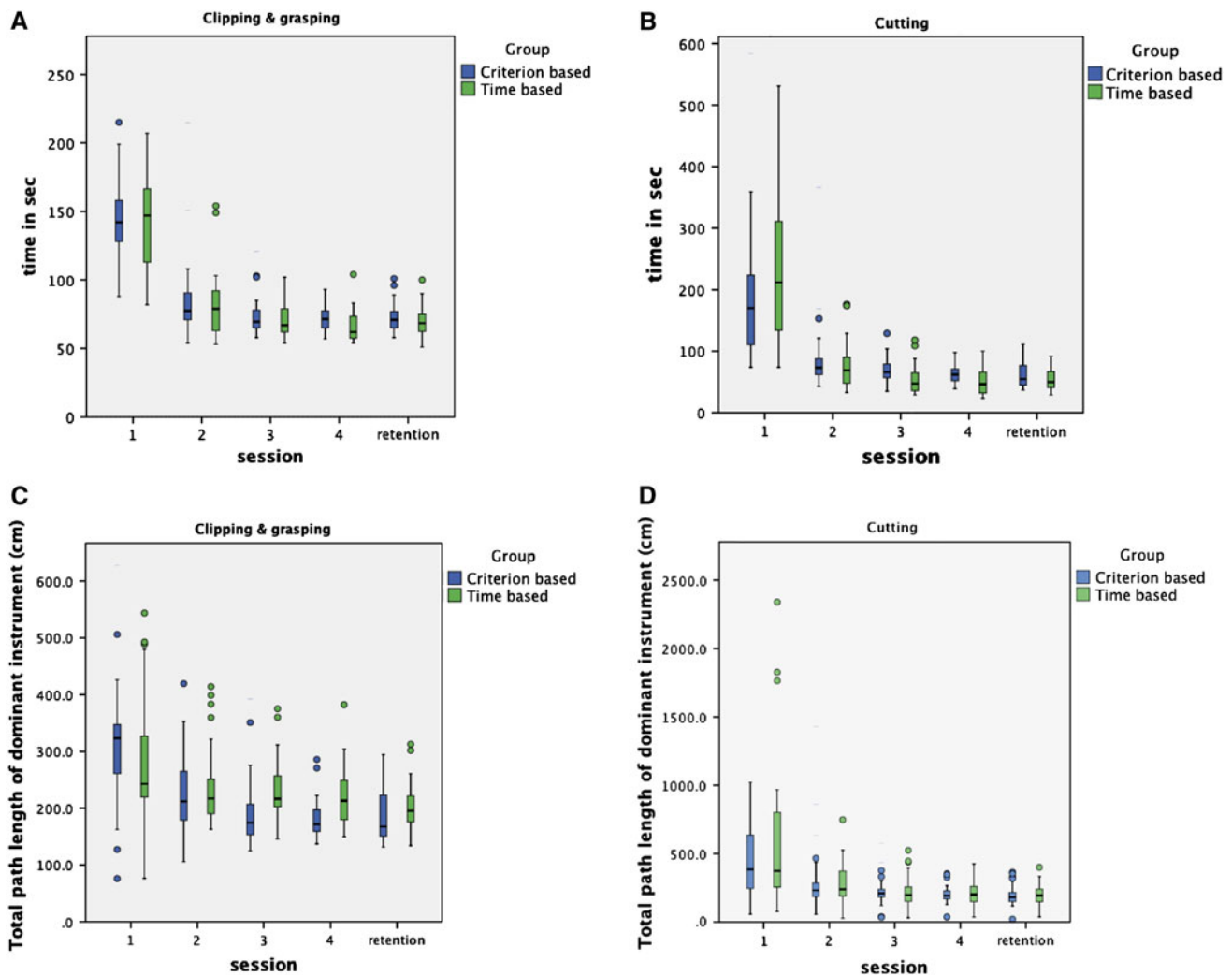


Fig. 2 Boxplot of LAP Mentor parameters: time for the **A** clipping and grasping task and **B** cutting task, and path length for the **C** clipping and grasping and **D** cutting task

respectively; $P = 0.001$) (Table 5). Retrospectively, the average number of repetitions needed to meet the criteria did not differ significantly between the groups (Table 4), although Group T was unaware of the criteria.

Discussion

In this study, we show that training time can be significantly reduced using criterion-based training instead of time-based training. Training results for the two training methods did not differ significant. We confirm that novices can extensively improve their skills in basic laparoscopy by training on the LAP Mentor. Both groups showed equal retention of skills. The skills learned on the LAP Mentor do transfer to a different laparoscopy simulator, the ProMIS.

Previous studies have shown advantages of criterion-based training in training outcome and in operating

performance [9, 10]. These studies did not describe training time benefits, because training time was fixed and equal in both groups. The absolute performance benefits of criterion-based training shown by Gauger et al. [10] were not found by our study. The fact that the criterion-based group (group C) in our study did not significantly outperform the time-based group (group T) can partly be explained by the significant differences in the total amount of repetitions and differences in the total amount of repetitions and the associated total training time in favour of group T. This was a direct consequence of the training protocol, forcing the participants in group T to continue training despite their performance.

Nevertheless it would be expected that group T, which trained with significantly more repetitions and longer duration, would outperform the other group in the post-test or retention level. This was not the case. The equal post-test performance in both groups despite the significantly fewer repetitions in group C can presumably be related to the

Table 2 Simulator scores Promis I ($n = 21$)

	Pre-test Mean (min–max)	Post-test Mean (min–max)	Retention Mean (min–max)	Pre-test–post-test <i>P</i> value*	Post-test–retention <i>P</i> value*
Sharp dissection					
Dominant instrument time (s)					
Group C	390 (217–644)	197 (125–293)	183 (124–341)	ns	ns
Group T	385 (182–697)	215 (106–468)	224 (94–414)	ns	ns
Dominant instrument path (mm)					
Group C	660 (333–1,146)	417 (246–631)	383 (252–508)	0.028	ns
Group T	686 (224–1,078)	486 (196–927)	441 (196–641)	0.016	ns
Object positioning					
Dominant instrument time (s)					
Group C	186 (137–266)	165 (95–302)	116 (72–152)	ns	ns
Group T	275 (113–493)	163 (112–293)	148 (91–213)	0.013	ns
Dominant instrument path (mm)					
Group C	506 (294–1,373)	683 (247–3,588)	344 (207–714)	ns	ns
Group T	596 (285–951)	432 (256–864)	399 (288–532)	0.041	ns

ns Not significant

* Within-group analysis, Wilcoxon signed rank

Table 3 Simulator scores ProMIS III ($n = 13$)

	Pre-test Mean (min–max)	Post-test Mean (min–max)	Retention Mean (min–max)	Pre-test–post-test <i>P</i> value*	Post-test–retention <i>P</i> value*
Sharp dissection					
Time (s)					
Group C	401 (276–511)	334 (144–611)	292 (181–416)	ns	ns
Group T	434 (255–615)	368 (173–624)	345 (215–592)	ns	ns
Economy of mov					
Group C	1,083 (717–1,464)	993 (443–1,984)	975 (609–1,491)	ns	ns
Group T	1,035 (679–1,383)	894 (463–1,603)	924 (582–1,460)	ns	ns
Object positioning					
Time (s)					
Group C	227 (169–273)	132 (109–185)	119 (92–173)	0.028	ns
Group T	189 (114–248)	129 (80–178)	123 (96–141)	0.028	ns
Economy of mov					
Group C	679 (359–962)	395 (270–585)	358 (255–490)	0.046	ns
Group T	504 (325–603)	355 (245–495)	348 (258–426)	0.028	ns

ns Not significant

* Within-group analysis, Wilcoxon signed rank

lower amount of overtraining in the criterion-based training. Some overtraining can be beneficial, although too much extra practice can lead to poor test performance [14]. In our study, group T had extensively extra practice; while criteria were reached after an average of 15 repetitions for the clipping and grasping task and 13 repetitions for cutting task, the average total repetitions were 46 and 45, respectively. The performances on the simulator did not improve significantly during that extra training. When using the criteria as optimal endpoint, there was approximately 200% overtraining in

repetitions. Group C did have some overtraining as well, because of the requirement to reach the criteria on every training day, however this was far less than for group T. It seems that identification of training criteria or benchmarks and a related training endpoint can reduce excessive over practice and corresponding unnecessary training time.

Another contradiction with previous research is that we did not find a significant difference in our study between the groups in terms of the number of repetitions needed to achieve the criteria, even though these were shown to the

Table 4 Number of repetitions required by the participants to achieve the criteria for the LAP Mentor tasks

	Criterion-based group mean (min–max)	Time-based group mean (min–max)	<i>P</i> value
Clipping and grasping	16.53 (6–43)	15.24 (4–32)	ns
Cutting	10.31 (2–25)	13.47 (2–63)	ns

Mann–Whitney *U* test, two-tailed

Note that participants in group T were not informed about the criteria
ns Not significant

Table 5 Number of task repetitions performed by the participants per day

	Criterion-based group mean (min– max)	Time-based group mean (min–max)	<i>P</i> value
Day 1	3.79 (1–7)	4.38 (2–14)	ns
Day 2	8.38 (2–23)	12.97 (6–28)	<0.001
Day 3	6.15 (2–17)	16.56 (10–25)	<0.001
Day 4	4.15 (2–11)	9.68 (3–14)	<0.001
Total	24.29 (14–60)	45.53 (28–78)	<0.001
Total training time (min)	73.71 (48.13–121.57)	120.17 (96.83–158.62)	<0.001

Mann–Whitney *U* test, two-tailed

ns Not significant

participants. In other studies [9, 10], when criteria were made known to the participants, the number of repetitions needed to meet the criteria did decrease.

There are two possible explanations for this contradiction. The first is that criteria were possibly set too easy, so that they were effortlessly reachable with or without known criteria. The second is the fact that the time-based group had knowledge of their results; they were equally aware of their performance and improved because of the feedback from the simulator after each exercise. This may have caused the participants in this group to train on improving their own scores, converting the training in some way to criterion-based training in which they set their own criteria.

Limitations

Due to two different test locations, we used two different ProMIS simulators. Because of different output settings, we could not perform a combined analysis.

Conclusions

The outcome of this study allows us to conclude that criterion-based training is more time efficient than time-based training in training of basic laparoscopic skills. It is recommended to develop future curricula as criterion-based. Therefore, one of the first steps in implementing new, or revising existing, curricula should consist of implementing criterion-based training with predefined criteria.

Acknowledgments The authors would like to thank Cecilia Morsika who helped with the inclusion of Greek participants.

Disclosures Authors W. M. Brinkman, S. N. Buzink, I. H. J. T. de Hingh, L. Alevizos and J. J. Jakimowicz have no conflicts of interest or financial ties to disclose.

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

1. Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ (2008) A systematic review of skills transfer after surgical simulation training. *Ann Surg* 248:166–179
2. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, Satava RM (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 236:458
3. Schijven M, Jakimowicz J, Broeders IAMJ, Tseng LNL (2005) The Eindhoven laparoscopic cholecystectomy training course—improving operating room performance using virtual reality training: results from the first EAES accredited virtual reality trainings curriculum. *Surg Endosc* 19:1220–1226
4. McDougall E, Corica F, Boker J, Sala L, Stolar G, Borin J, Chu F, Clayman R (2006) Construct validity testing of a laparoscopic surgical simulator. *J Am Coll Surg* 202:779–787
5. Zhang A, Hunerbein M, Dai Y, Schlag P, Beller S (2008) Construct validity testing of a laparoscopic surgery simulator (Lap Mentor®). *Surg Endosc* 22:1440–1444
6. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time (2003). *Official Journal L* 299, 18/11/2003 P 0009-0019
7. Verdaasdonk E, Stassen L, van Wijk R, Dankelman J (2007) The influence of different training schedules on the learning of psychomotor skills for endoscopic surgery. *Surg Endosc* 21:214–219
8. Mackay S, Morgan P, Datta V, Chang A, Darzi A (2002) Practice distribution in procedural skills training. *Surg Endosc* 16:957–961
9. Ahlberg G, Enochsson L, Gallagher A, Hedman L, Hogman C, McClusky D III, Ramel S, Smith C, Arvidsson D (2007) Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 193:797–804
10. Gauger P, Hauge L, Andreatta P, Hamstra S, Hillard M, Arble E, Kasten S, Mullan P, Cederna P, Minter R (2010) Laparoscopic

- simulation training with proficiency targets improves practice and performance of novice surgeons. *Am J Surg* 199:72–80
11. Stefanidis D, Acker CE, Greene FL (2010) Performance goals on simulators boost resident motivation and skills laboratory attendance. *J Surg Educ* 67:66–70
 12. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J (2003) Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg* 185:146–149
 13. Heinrichs W, Lukoff B, Youngblood P, Dev P, Shavelson R, Hasson H, Satava R, McDougall E, Wetter P (2007) Criterion-based training with surgical simulators: proficiency of experienced surgeons. *JSL* 11:273–302
 14. Magill RA (2011) *Motor learning and control*. McGraw-Hill, New York