

## Suturing training in Augmented Reality: gaining proficiency in suturing skills faster

S. M. B. I. Botden · I. H. J. T. de Hingh ·  
J. J. Jakimowicz

Received: 10 July 2008 / Accepted: 7 October 2008 / Published online: 6 December 2008  
© The Author(s) 2008. This article is published with open access at Springerlink.com

### Abstract

**Background** Providing informative feedback and setting goals tends to motivate trainees to practice more extensively. Augmented Reality simulators retain the benefit of realistic haptic feedback and additionally generate objective assessment and informative feedback during the training. This study researched the performance curve of the adapted suturing module on the ProMIS Augmented Reality simulator.

**Methods** Eighteen novice participants were pretrained on the MIST-VR to become acquainted with laparoscopy. Subsequently, they practiced 16 knots on the suturing module, of which the assessment scores were recorded to evaluate the gain in laparoscopic suturing skills. The scoring of the assessment method was calculated from the “time spent in the correct area” during the knot tying and the quality of the knot. Both the baseline knot and the knot at the top of the performance curve were assessed by two independent objective observers, by means of a standardized evaluation form, to objectify the gain in suturing skills.

**Results** There was a statistically significant difference between the scores of the second knot (mean 72.59, standard deviation (SD) 16.28) and the top of the performance curve (mean 95.82, SD 3.05;  $p < 0.001$ , paired t-test). The scoring of the objective observers also differed significantly (mean 11.83 and 22.11, respectively; SD 3.37 and 3.89, respectively;  $p < 0.001$ ) (interobserver reliability Cronbach’s alpha = 0.96). The median amount of repetitions to reach the top of the performance curve was eight, which also showed significant differences between both the assessment score (mean 88.14, SD 13.53,  $p < 0.001$ ) and scoring of the objective observers of the second knot (mean 20.51, SD 4.14;  $p < 0.001$ ).

**Conclusions** This adapted suturing module on the ProMIS Augmented Reality laparoscopic simulator is a potent tool for gaining laparoscopic suturing skills.

**Keywords** Laparoscopy · Simulation · Training · Performance curve · Laparoscopic suturing

---

S. M. B. I. Botden (✉) · I. H. J. T. de Hingh · J. J. Jakimowicz  
Department of Surgery, Catharina Hospital, Michelangelolaan 2,  
5623 EJ Eindhoven, The Netherlands  
e-mail: Sanne.Botden@cze.nl; Sanne\_Botden@hotmail.com

I. H. J. T. de Hingh  
e-mail: Ignace.d.Hingh@cze.nl

S. M. B. I. Botden · J. J. Jakimowicz  
Faculty of Industrial Design Engineering, Delft University  
of Technology, Delft, The Netherlands

J. J. Jakimowicz  
e-mail: Jack.jakimowicz@cze.nl

The use of simulation in surgical training curricula is becoming more widely accepted and most Virtual and Augmented Reality simulators are able to provide objective assessment and feedback. Objective assessment of performance is fundamental to provide formative feedback during training, allowing for continuous skill refinement [1]. Providing feedback and setting goals tends to motivate trainees to practice their skills more extensively compared with a self-directed group [2].

The development of objective measures of operative skill is important to confirm the role of simulators in laparoscopic surgery [3]. To be an effective training tool, the simulator has to provide metrics that are meaningful and

informative to the trainee. Professional organizations have recently recognized the need to assess surgical performance objectively [3].

Haptic feedback is fundamental for good laparoscopic training, in particular for laparoscopic suturing [4–6]. Laparoscopic training with haptic feedback results in significantly improved skills transfer to the trainee compared with training without haptic feedback [3]. Especially for laparoscopic suturing skills, it is important that the participants have tactile feeling of what they are doing during the procedure. In general, it is assumed that realistic simulations with haptic feedback result in better training effects and better transfer to the clinical setting [7]. Augmented Reality provides realistic haptic feedback because of the hybrid mannequin environment in which the trainee is working, with real laparoscopic instruments and materials.

To become proficient in laparoscopic skills or procedures, surgeons in training must experience their own learning curve to gain proficiency. Improvement in skills tends to be more rapid during the first part of the training and will taper off over time until a steady state has been reached [8]. In this study we examined the performance curve on suturing skills of trainees with no previous laparoscopic suturing experience to research whether the suturing module of the ProMIS v2.0 Augmented Reality simulator is proficient for training of suturing skills to surgical residents.

## Methods

### Subjects

In total 18 novice participants completed the 2-day training sessions on the suturing module of the ProMIS V2.0. All participants were medical students during their clinical rotations or surgical residents. None of the participants had previous clinical laparoscopic experience or any laparoscopic suturing experience. An informed consent was signed by all participants to state that they voluntarily participated in this study.

### Equipment

#### *ProMIS V2.0*

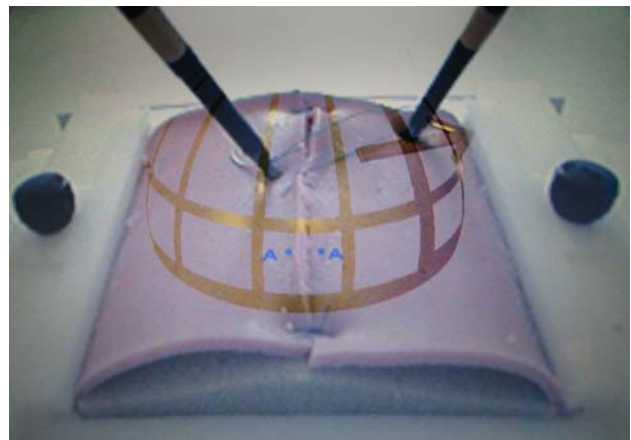
In this study we used the ProMIS v2.0 Augmented Reality (AR) simulator (Haptica, Dublin, Ireland). The laparoscopic interface was a torso-shaped mannequin (29" L × 20" W × 9" D) with a skin-colored cover, which is connected to a notebook (Dell, XPS M1710). The mannequin contained three separate camera tracking systems, arranged to identify any instrument inside the simulator from three different

angles. The camera tracking systems captured instrument motion with Cartesian coordinates in the x, y, and z planes at the average rate of 30 frames per second (fps). The distal end of the laparoscopic instrument shaft was covered with two pieces of yellow electrical tape to serve as a reference point for the camera tracking system; therefore, it accepted a broad range of instrument types. Instrument movement was recorded and stored in distinct sections, based on the time the tips of the instrument was detected until removed from the mannequin. The notebook was positioned so that the participant had the screen placed just below eye level and the mannequin was placed at a standard ergonomic height for performing the laparoscopic tasks.

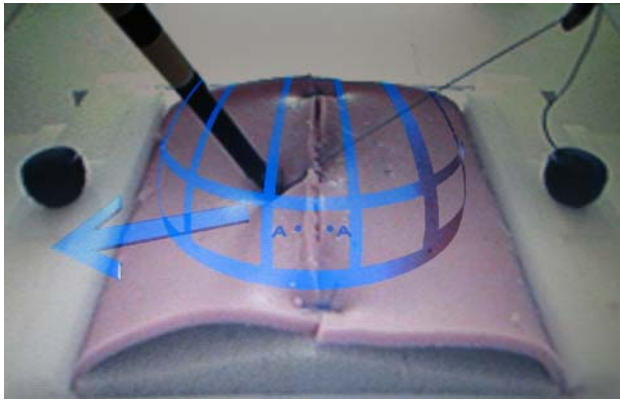
The simulator recorded time, path length, and smoothness of movement (through changes in instrument velocity and changes in direction), during each separate task within the training module. After completion of the task, ProMIS provided statistics on the screen. In addition, a full video and virtual playback of the trainee's performance were saved. The suturing pads for the suture and knot tying task were placed inside the mannequin.

### *Suturing module*

For the suturing module used in this study, an adapted assessment method was developed, which calculated the assessment scores based on the time spent in the correct area (dome) and the quality (strength) of the tied knot. The dome itself was only for the path of throwing the thread over the needle holder. When the thread had to be pulled tight, the instrument had to come out of the dome at a certain point and angle, which was calculated from the ideal path of the experts and guided by an arrow on the screen (Fig. 1). It was important that the trainee only pulled on the proper side of the thread and only with the proper



**Fig. 1** Dome of the adapted suturing module is visualized on the screen as guidance and assessment method during the suturing task. When the instruments are inside the proper area, the dome is translucent



**Fig. 2** Dome and arrow during the suturing task to guide the trainee in the proper direction when pulling the knot tight. The dome turns bright blue when pulling the knot tight in the wrong direction

hand in the correct angle to create a surgeon's knot. At the end of the task the assessment score of the performance was shown by means of a percentage per step that was spent in the correct area (Fig. 2). The quality (strength) of the knot was tested by cutting the suture out of the suturing pad and pulling at the cut ends with a tension meter. This showed whether the knot would slip or break when pulled with at least 25 N, which a correct surgeon's knot should be able to endure [9].

During training, the 26173 KL and 26173 KAL KOH macro needle holders (Karl Storz, Tutlingen, Germany) with Covidien Polysorb 3-0 suturing needle and thread were used.

#### *MIST-VR*

To practice the basic skills of laparoscopic surgery and become acquainted with the fulcrum effect, the participants practiced four tasks on the well-validated MIST-VR laparoscopic simulator: Acquire Place, Withdrawal, Transfer Place, and Traversal [4, 10–13].

#### Performance evaluation form

The assessment of the suturing skills of the baseline knot and the final knot were performed by using a standard evaluation form, which consisted of seven items scored on five-point Likert scale: 1) positioning needle in needle holder; 2) running needle through suturing pad; 3) taking proper bites of the suturing pad while performing the suture; 4) throwing thread around needle holder; 5) pulling tight the thread in proper direction; 6) tying a correct surgical knot; and 7) global evaluation of performance. For this study we used the summation of these seven items to compare the performances. Both expert observers have performed more than 500 clinical laparoscopic procedures and have extensive laparoscopic suturing experience, using the same suturing and knot-tying technique.

#### Protocol

Eighteen novice participants were pretrained on the MIST-VR with four basic tasks before starting the suturing sessions to become acquainted with the laparoscopic basics and the fulcrum effect. Thereafter, they watched a demonstration video of a laparoscopic surgeon's knot. During the training, step-by-step instruction videos were used to guide the trainee during practice, whereas no additional verbal feedback was given on the suturing skills.

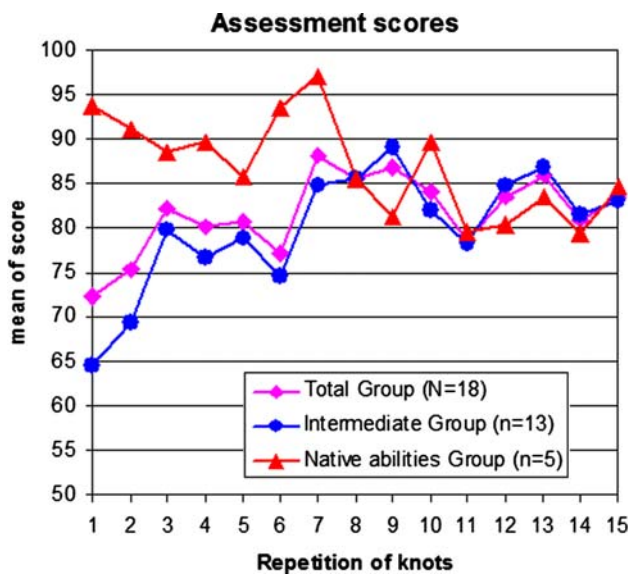
Then, the participants practiced two knots on the adapted suturing module on the ProMIS V2.0, of which the second one was assessed as the baseline score. The first knot was not used as the baseline to avoid bias in the results, caused by unfamiliarity with both the simulator, instruments, and module; therefore, the second run of the suturing task was used as a baseline knot. The second (baseline) knot was assessed by two independent objective observers, by means of the standard evaluation form. The participants practiced another 14 knots on the suturing module in 2 days: in total 7 knots on the first day and 9 on the second day. After the training session, the recorded knots at the top of the individual performance curve and at the overall performance curve (knot 7) were assessed by the objective observers. During the training, the assessment scores of the adapted assessment module were gained to evaluate the gain in suturing skills and to visualize the learning effect of this suturing module.

#### Statistics

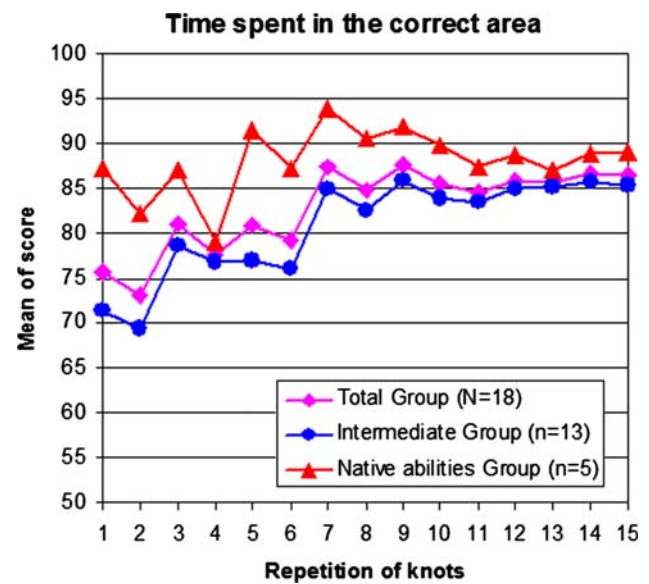
All data was processed and analyzed by using SPSS 13.0 (SPSS, Inc., Chicago, IL). Differences between the performance scores during the training were calculated with ANOVA. The differences between the performance scores at the baseline and at the top of the performance curve were calculated with the paired *t* test. Interobserver reliability was calculated with Cronbach's alpha.  $p < 0.05$  was considered a significant difference.

#### Results

According to the assessment method, a median of eight repetitions were necessary to reach the top of the performance curve (Fig. 3); This is knot 7 in the tables and figures, because the first knot was not used in the results to avoid bias. The individual top of the performance curve was compared with the baseline scores, which showed a significant difference in the assessment score (mean 72.59 vs. 95.82,  $p < 0.001$ ). There also was a significant difference between the assessment scores of the second (baseline) knot and knot 7 (mean 72.59 and 88.14, respectively;  $p = 0.001$ ; Table 1).



**Fig. 3** Assessment scores during the suturing training improved significantly:  $p = 0.008$  for the total group, and  $p < 0.001$  for the intermediate group (ANOVA). Participants with native abilities did not show significant differences in their performance ( $p = 0.602$ , ANOVA)



**Fig. 4** “Time spent in the correct area” during the suturing training improved significantly:  $p < 0.001$  for both the total group and intermediate group (ANOVA) but not for the native abilities group,  $p = 0.209$  (ANOVA)

The scores of “time spent in the correct area” reached the plateau phase at knot 7 (Fig. 4) and showed significant improvement during the training ( $p < 0.001$ , ANOVA). The scores of the “strength of the knot” tended to vary during the training session and did not show a significant difference ( $p = 0.479$ , ANOVA), with a dip in the scores at knots 6 and 11 (Fig. 5). Although this was not a primary assessment parameter, there was a significant decrease in the time to complete the knot during the training ( $p < 0.001$ , ANOVA).

The scoring of the objective observers differed significantly for both the knot at the top of the individual performance curve and knot 7 compared with the baseline knot (mean 11.83, 22.11, and 20.51, respectively;  $p < 0.001$ ; Table 1), with an interobserver reliability Cronbach’s alpha of 0.96.

The scores of the individual assessment parameters “time spent in the correct area” and “strength of knot” also showed significant differences between the baseline knot and both the knot at the top of the average and individual

performance curve ( $p = 0.003$ ,  $p = 0.004$ , and  $p < 0.001$ , respectively; Table 1).

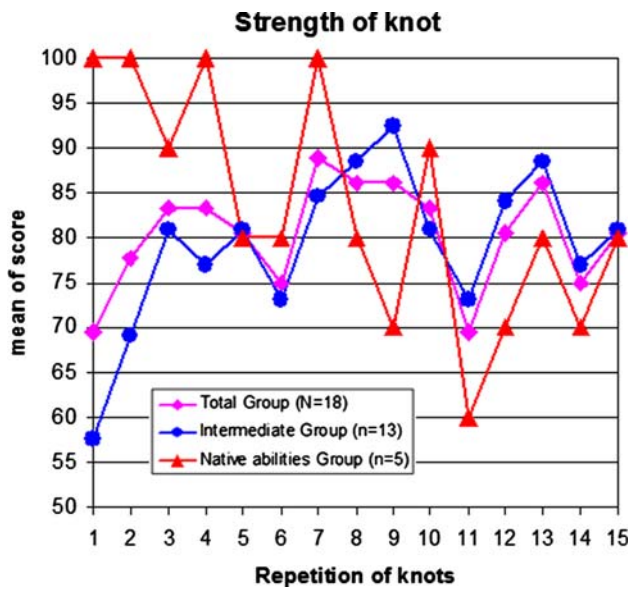
Five participants started the training session with high assessment scores and were able to perform a correct surgeon’s knot while staying in the correct area for the major part of the performance. Their assessment scores did not improve significantly during the training ( $p = 0.602$ , ANOVA; Fig. 3) nor comparing the baseline knot with the best performance ( $p = 0.08$ ; Table 2). “Time spent in the correct area” did not show significant improvement during the training session (Fig. 4) and neither did “strength of the knot,” which decreased during the progress of the training (Fig. 5). The average group ( $n = 13$ ) significantly improved both the assessment score and “time spent in the correct area” during training ( $p < 0.001$ ; Figs. 3 and 4, Table 2). The strength of the knot did not improve significantly during the training process, because there are a few dips in the curve (Fig. 5).

**Table 1** Gaining proficiency in suturing skills

	Mean (standard deviation)			p-value	
	Baseline knot	Knot 7	Knot on top of individual curve	Baseline vs. knot 7	Baseline vs. knot on top
Assessment score	72.59 (16.28)	88.14 (13.53)	95.82 (3.05)	<0.001	<0.001
Time spent in correct area	76.3 (15.06)	87.39 (12.91)	91.64 (6.11)	0.003	<0.001
Strength of knot	69.44 (25.08)	88.89 (21.39)	100 (0)	0.004	<0.001
Objective observer score	11.83 (3.37)	20.51 (4.14)	22.11 (3.89)	<0.001	<0.001

Differences were calculated by using the paired  $t$  test;  $p < 0.05$  was considered a significant difference





**Fig. 5** “Strength of the knot” did not show significant improvement during the training for the total, intermediate, and native abilities groups ( $p = 0.479, 0.105, \text{ and } 0.104$ , respectively, ANOVA)

Before the participants started the suturing training, they were all pretrained on the MST-VR simulator with four basic tasks. The performance scores from these tasks were compared with the scores of the baseline knot to determine whether there were correlations between the native psychomotor abilities and the native suture and knot-tying skills. No significant correlations were found between the basic and suturing skills (Table 3). However, there was a correlation in the performance on the MIST-VR and the time to complete the baseline suture task, but because time is not a primary assessment measurement, no conclusions can be made. Moreover, a significant negative correlation was calculated between the scores of the MIST-VR and the “time spent in the correct area” during the suturing.

Two participants had major difficulties with suturing and knot-tying skills on the first training day: one was not able to finish the baseline knot and was excluded from the results, and the other was able to tie a correct surgeon’s knot while staying in the correct area for a respectable part of the performance on the second day.

**Table 3** Correlation between basic skills and suturing

	Spearman’s rho	$p$
Assessment score	−0.185	0.463
Time spent in the correct area	−0.479	0.044
Strength of the knot	0.099	0.696
Time to complete the performance	0.637	0.004

Correlation between the performance on the MIST-VR basic laparoscopic tasks and the laparoscopic suturing performance at the baseline knot, calculated with Spearman’s rho

$p < 0.05$  was considered a significant correlation

**Discussion**

The term feedback refers to the return of performance-related information to the performer and can be divided into two major categories: intrinsic feedback and extrinsic feedback [14, 15]. Intrinsic feedback consists of performance-related information available directly to the sensory system, such as haptic feedback and visual cues of the instrument movement during the task [14, 15]. Extrinsic feedback is performance-related information provided by an external source and has two important roles: 1) to facilitate achievement of the performance goal, by providing information about the degree of success thus far, and about the various components involved in achieving that performance goal; and 2) extrinsic feedback should motivate the trainee to continue to strive toward the achievement of that goal [14, 15]. To motivate trainees to practice their skills, this extrinsic feedback should be meaningful and informative [15]. Previous studies have shown that although extrinsic feedback can provide insight into actions and consequences of the actions, they also can inhibit intrinsic learning strategies and the development of problem-solving abilities [14].

Stefanidis et al. [14] suggested a trend toward faster achievement of simulator proficiency with the incorporation of frequent video tutorial viewing. Providing video demonstrations before and during training, as used in this study, has been shown to lead to superior training [14].

It is important that trainees understand the extrinsic feedback that is provided, to translate this in improvement

**Table 2** Mean assessment score

	Mean (standard deviation)			$p$
	Baseline knot	Knot 7	Knot in top of individual curve	
Average abilities ( $n = 13$ )	64.51 (10.85)	84.75 (14.58)	95.25 (3.27)	Baseline vs. knot 7: 0.001 Baseline vs. top: <0.001
Native abilities ( $n = 5$ )	93.6 (2.6)	96.94 (2.39)	*	0.08

Differences were calculated by using the paired  $t$  test;  $p < 0.05$  was considered a significant difference

\* Knot 7 scores were the highest on the individual performance gain curves of the group with the native abilities

of the skills. Often-used parameters, such as time and path length, are not readily transferable to informative feedback to the trainee, and cannot be used to effectively improve laparoscopic suturing skills.

In this study the strength of the knot was tested, which provided informative extrinsic feedback of the performance, as a proper tied knot is the performance goal in laparoscopic suturing training. “Time spent in the correct area” was another measurement used to calculate the performance score and was shown on the screen after each performance. The trainee was able to see the most problematic part of the knot tying and what could be done to achieve an optimal knot.

There was a dip in performance at knots 6 and 11 (Fig. 3). Knot 6 was the last knot on the first day of the training, during which the participants complained of tiredness and were no longer focused. Knot 11 was the fifth knot of the second training day. On the second day, knot tying improved more than expected for most participants, which resulted in losing focus and concentration. Practicing too intensely in one day may exhaust a trainee, thus negatively influencing performance. This could potentially cause the trainee loss of motivation and negative extrinsic feedback. Therefore, the recommendation to spread the laparoscopic training over several days or reduce the amount of sutures in one session seems to be justified. The first knot on the second day (knot 7) was the best knot on average, presumably because the participants had the opportunity to recapitulate both the intrinsic and extrinsic feedback overnight and regain their concentration.

It has been indicated that learning can be enhanced when trainees have the opportunity to practice with progressively increasing levels of difficulty [10, 16, 17]. In this study all participants remained in the “beginner level” mode during the training. However, three difficulty levels are available, which have been developed to motivate trainees to practice their skills extensively until they have reached the advanced skills level. For the participants with a native ability in laparoscopic suturing, an “advanced level” mode could be an option to increase their motivation. The performance scores in the “beginner level” mode decreased during training; if the participants had trained in the more challenging “intermediate level” mode, perhaps they would have remained focused to perform with their best capabilities. Existence of a relationship between cognitive abilities and skills acquisition in the early phase of learning new surgical skills has been debated, but these correlations seems to decline when the procedure becomes routine [18, 19]. This could be another explanation for the high performance scores during the first runs of the task and the decrease in the scores at the end of the training. However, pretraining on the MIST-VR does not show a correlation with the scores of the baseline knot, which

should otherwise visualize the native psychomotor abilities of the trainee.

## Conclusions

This adapted suturing module on the ProMIS Augmented Reality laparoscopic simulator is a potent tool for the training of laparoscopic suturing skills to surgical residents. The trainees in this study needed only seven repetitions on average to reach the top of the performance curve. There were statistically significant differences for both the scoring of this assessment method and the scoring by the objective observers when comparing the baseline knot with the top of the performance curve.

**Acknowledgements** The authors thank Cees Schot for technical support. This study was partly funded by The Scientific Foundation of the Catharina Hospital Eindhoven, The Netherlands and Covidien, The Netherlands. This study was performed by objective researchers who have no attachments with the industry.

**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. Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG (2004) Proving the value of simulation in laparoscopic surgery. *Ann Surg* 240:518–528
2. Gonzales R, Bowers SP, Smith CD, Ramshaw BJ (2004) Does setting specific goals and providing feedback during training resulting better acquisition of laparoscopic skills? *Am Surg* 70:35–39
3. Aggarwal R, Moorthy K, Darzi A (2004) Laparoscopic skills training and assessment. *Br J Surg* 91:1549–1558
4. 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* 91:146–150
5. 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* 236:458–464
6. Van Sickle KR (2005) Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg Endosc* 19:1227–1231
7. Strom P, Hedman L, Sarna L, Kjellin A, Wredmark T, Fellander-Tsai L (2006) Early exposure to haptic feedback enhances performance in surgical simulator training: a prospective randomized crossover study in surgical residents. *Surg Endosc* 20:1383–1388
8. Schijven MP, Jakimowicz JJ (2004) The learning curve on the Xitact LS 500 laparoscopy simulator: profiles of performance. *Surg Endosc* 18:121–127
9. Hanna GB, Frank TG, Cuschieri A (1997) Objective assessment of endoscopic knot quality. *Am J Surg* 174:410–413

10. 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* 191:128–133
11. Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D (2006) Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? *Surg Endosc* 20:104–112
12. Van Sickle KR, Ritter EM, McClusky DA 3rd, Lederman A, Baghai M, Gallagher AG, Smith CD (2007) Attempted establishment of proficiency levels for laparoscopic performance on a national scale using simulation: the results from the 2004 SAGES Minimally Invasive Surgical Trainer-Virtual Reality (MIST-VR) Learning Center Study. *Surg Endosc* 21:5–10
13. 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* 19:1523–1532
14. Stefanidis D, Korndorffer JR, Heniford BT, Scott DJ (2007) Limited feedback and video tutorials optimize learning and resource utilization during laparoscopic simulator training. *Surgery* 142:202–206
15. Porte MC, Xeroulis G, Reznick RK, Dubrowski A (2007) Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *Am J Surg* 193:105–110
16. Issenberg BS, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ (2005) Features and use of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teacher* 27:10–28
17. Ali MR, Mowery Y, Kaplan B, DeMaria EJ (2002) Training the novice in laparoscopy. More challenge is better. *Surg Endosc* 16:1732–1736
18. Keehner MM, Tendick F, Meng MV, Anwar HP, Hegarty M, Stoller ML, Duh QY (2004) Spatial ability, experience, and skill in laparoscopic surgery. *Am J Surg* 188:71–75
19. Keehner MM, Lippa Y, Montello DR, Tendick F, Hegarty M (2006) Learning a spatial skill for surgery: how the contribution of abilities change with practice. *Appl Cogn Psych* 20:487–503