Surg Endosc (2005) 19: 1220–1226
DOI: 10.1007/s00464-004-2240-1

SURGICAL ENDOSCOPY
and Other Interventional Techniques

The Eindhoven laparoscopic cholecystectomy training course—improving operating room performance using virtual reality training

Results from the first E.A.E.S. accredited virtual reality trainings curriculum

M. P. Schijven, J. J. Jakimowicz, I. A. M. J. Broeders, L. N. L. Tseng

IJsselland Hospital, 2900 AR Capelle a/d IJssel, Post Office Box 696, The Netherlands

Received: 30 September 2004/Accepted: 10 May 2005/Online publication: 28 July 2005

Abstract
Background: This study was undertaken to investigate operating room performance of surgical residents, after participating in the Eindhoven virtual reality laparoscopic cholecystectomy training course. This course is the first formal surgical resident trainings course, using a variety of complementary virtual reality (VR) skills training simulation in order to prepare surgical residents for their first laparoscopic cholecystectomy. The course was granted EAES certification.
Methods: The four-day course is based on multimedia and multimodality approach. A variety of increasingly difficult simulation training sessions, next to intimate focus-group "knowledge sessions" are included. Both basic and procedural VR simulation is featured, using MIST-VR and the Xitacts' LapChol simulation software. The operating room performance of twelve surgical residents who participated in the course and twelve case-control counterparts were compared. The case-control grup was matched for clinical number laparoscopic cholecystectomy performance (maximum of 4 procedures). Two observers analyzed a randomly mixed videotape, featuring the part of the "clip-and-cut" procedure of the laparoscopic cholecystectomy, and were blinded for participants' group status. Structured questionnaires including multiple observation scales were used to assess performance.
Results: Residents of both the experimental and control group did not differ in demographic parameters, except for number of laparoscopic cholecystectomies in favor of the control group (*p*-value 0.008). Both observers judge the experimental group to perform significantly better (*p*-value 0.004 and 0.013). Experimental group residents valued their course highly in terms of their laparoscopic surgical skills improvement and the use of VR simulators in the surgical curiculum.

Conclusions: The Eindhoven Virtual Reality laparoscopic cholecyctectomy training course improves surgical skill in the operating room above the level of residents trained by a variety of other training methods.

Key words: Laparoscopic cholecystectomy — Virtual reality training — Virtual reality simulation — Surgical training — Skills laboratory

For more than a decade now, laparoscopic cholecystectomy has been the treatment of choice for uncomplicated cholelithiasis. The introduction of laparoscopic cholecystectomy and the rapid evolution of minimal access surgery (MAS) have challenged the classical apprenticeship model for surgical training. In fact, it is known that with the introduction of MAS, complications are bound to arise during surgeons' early experience with the particular procedure [16]. Currently, because of shortened surgical curricula, stress on working hours, and medico-legal issues, MAS training needs to be intensified to ensure safe and high-quality patient care. The traditional apprenticeship model for teaching is under debate, because it is costly in terms of time and resources and unlikely to provide adequate training for the skills needed [2, 10]. In fact, such skills cannot be extrapolated directly from those acquired in open surgery because MAS requires quite distinct psychomotor abilities, hand–eye coordination, and different skills. Moreover, the surgeon engaging in MAS has to overcome the hurdles of the two-dimensional videoscopic surgical interface, the restricted degrees of freedom of movement for the MAS instruments, and different haptic sensations.

Correspondence to: M. P. Schijven

Virtual reality (VR) surgical simulators provide excellent opportunities for enhancing psychomotor skill and training MAS procedures in a safe environment. Such systems allow repeated, unbiased practice of a standardized task, provided they are well validated in order to be accepted by the surgical community. In the literature, few studies have assessed the transfer of skill using VR simulation to actual operation room performance. The majority of these studies indeed have showed improved performance [4–6, 14], although some have failed to do so. [1] These studies all feature basic psychomotor skills trainers such as the MIST-VR (Mentice Medical Simulation, Gothenburg, Sweden), and LapSim (Surgical Science Ltd, Gothenburg, Sweden).

Recent developments in VR simulation have put simulators beyond the level of basic computerized psychomotor skills trainers. The newer generation of VR simulators and simulations are in fact able to mimic a MAS surgical procedure convincingly. No studies have investigated operating room performance using such a second-generation VR simulation.

The aim of the current study was to investigate the operation performance of surgical residents after their participating in the Eindhoven virtual reality laparoscopic cholecystectomy training course. This course uses the Xitact LS500 as a second-generation open platform VR simulator. The procedure of the laparoscopic cholecystectomy is featured. Participants' operating room performance was compared with the operating room performance of a matched control group.

Materials and methods

Inclusion

The study investigated surgeons-in-training working in 1 of 11 participating teaching hospitals. Participants were recruited between April 2003 and March 2004. All were surgeons-in-training and novices in laparoscopic cholecystectomy.

Twelve surgeons participating in the Eindhoven course constituted the experimental group, and 12 other surgeons-in-training constituted the case–control group. All the participants, both in the experimental and the case–control group, attended a Basic Surgical Skills Course before participating in this study. None of the participants engaged in a specific cholecystectomy skills training course. Inclusion in either group was restricted to participants who had performed no more than four laparoscopic cholecystectomies. Patients selected for the course were patients considered American Society of Anesthesiology (ASA) class 1, with a medical history of uncomplicated cholelithiasis and no previous abdominal complaints or surgery.

Course

The surgeons in the experimental group participated in the 4-day virtual reality laparoscopic cholecystectomy training course at the Catharina Hospital Eindhoven, The Netherlands. The participants were insured as active surgical residents for the duration of the course. The course was set up to incorporate a variety of teaching elements relevant to the procedure under study. Videos featuring the use of laparoscopic cholecystectomy for both uncomplicated and complicated cases were discussed alongside a variety of oral presentations. Table sessions on the interactive "transfer of laparoscopic cholecystectomy" knowledge were incorporated, and laparoscopic instruments and stack (e.g., camera, light source) needed for performance of laparoscopic cholecystectomy were explained and displayed.

Core elements of the course were repetitive training sessions on a variety of VR software simulations using the open Xitact LS500 laparoscopy simulator platform (Xitact SA, Morges, Switzerland). Increasingly difficult levels of both basic psychomotor VR simulation (MIST-VR) and procedural laparoscopic cholecystectomy simulation, including the clip-and-cut, navigation, and dissection modules (Xitact), were featured. Thus, an integrated, procedure-specific, multimodality VR training curriculum was presented to the participants.

On days 2 and 3 of the course, the residents attended the operating room (OR) in conjunction with their VR training sessions to act as either an assistant surgeon or a camera assistant helping with a laparoscopic cholecystectomy being performed by an expert laparoscopic surgeon. On day 4 of the course, the participants performed a full laparoscopic cholecystectomy themselves under the close supervision of the expert surgeon. The procedure was videotaped.

The course was supervised by a faculty member assigned by the Scientific, Educational and Programme Committee of the European Association for Endoscopic Surgery (EAES). Subsequently, the course accredited formal EAES certification.

Assessment

The participants in both research groups performed the laparoscopic cholecystectomy procedure under expert surgical supervision. Only the clip-and-cut part of the laparoscopic cholecystectomy (e.g., the clipping and cutting of the cystic artery and cystic duct was the object of study. This part was chosen for multiple reasons. First, it is likely to be one of the most essential stages of the laparoscopic cholecystectomy procedure that must be performed safely to avoid possible damage to the common bile duct.

Second, by selective assessment of the clip-and-cut scenery, technical skills can be assessed quite independently of other factors influencing outcome variation. Third, it is likely to be a good predictor of overall performance. Finally, by monitoring this specific part of the procedure, assessment is facilitated because the outcome assessment scale can be properly standardized.

The procedure was assessed starting from the moment the laparoscopic clip applier was introduced and ending at the moment the laparoscopic scissors were removed from the operative field. Video fragments from both the experimental and control group residents were evaluated by two reviewers, both laparoscopic engaged surgeons from different academic training hospitals, independently of each other. The participant's video fragments were mixed in random order before being copied to the reviewers videotape. In both the experimental and the control group, two recordings were excluded from the analysis as a result of technical recording failure. Statistical analysis assessing the data of the remaining 20 procedures was performed using the SpSS Version 10.0 software package (SpSS, Chicago, IL, USA).

Outcome parameters

A structured questionnaire using a 5-point Likert rating scale ranging from 0 (completely disagree) to 5, (completely agree) was used for assessment. The parameters of interest were "fluency," which was operationalized by the statement "Pattern of movement is fluent, precise, and efficient (few unnecessary/random movements)," and "carefulness," which was operationalized by the statement "The tissue is treated with respect during the procedure, without visible excessive force, traction, or resulting in injury."

A specific rating scale termed "sumscore" was designed to focus on the phase of the cholecystectomy clip-and-cut procedure with this scale, performance is judged according to an integration of psychomotor skills, procedural knowledge of anatomy, and decision making. This sumscore was constructed primarily according to the metrics used as a framework for the assessment of Xitact's clip-and-cut simulation [10].

The final outcome parameter, "judgement" was operationalized by the following question: "For the procedure of clipping and cutting of the cystic duct and cystic artery by this surgeon, I would grade him or her a —" (number in the range of 1 to 10, carried to one decimal). Time to complete the clip-and-cut procedure was recorded.

Table 1. Demographics by group (experimental vs control)

	Sex	Age	Year of training	Specialty	Number of laparoscopic cholecystectomies	
					Performed	Assisted
Mann-Whitney U	40,000	43,500	34,000	31,500	14,500	24,500
Asymptomatic significance (two-tailed)	0.342	0.900	0.317	0.201	0.008[a]	0.083

[a] Significant difference $p \leq 0.05$

Results

Demographics

The mean age of the participants in both groups was 31 years. All the participants were right-handed. There were 4 females in the experimental group, whereas there were 2 females in the control group. The mean years of training was 1.8 for both groups.

All the participants in both the experimental and the control groups were either in training for general surgery or in the first two mandatory general surgery training years while in training for another surgical subspecialty. Half (50%) of participants in the experimental group were in training to become (general) surgeons, in contrast to 70% of the participants in the control group. Two participants in the control group had engaged in an animal gallbladder training course, and one had taken a nonanimal one. The participants in the experimental group were not previously engaged in specific cholecystectomy training courses.

Of the demographic variables, only the number of laparoscopic cholecystectomies performed before participation in the study differed significantly (Table 1; $p = 0.008$, Mann–Whitney U test. The mean number of laparoscopic cholecystectomies performed in the experimental group was 0.3 (range, 0–1), whereas the mean number of procedures performed in the control group was 1.8 (range, 0–1), whereas the mean number of procedures performed in the control group was 1.8 (range, 0–3).

Course satisfaction

Figure 1 depicts the course satisfaction 95% confidence intervals according to course participants, opinion, in reference to the satisfaction statements 1-10 presented in (Table 2).

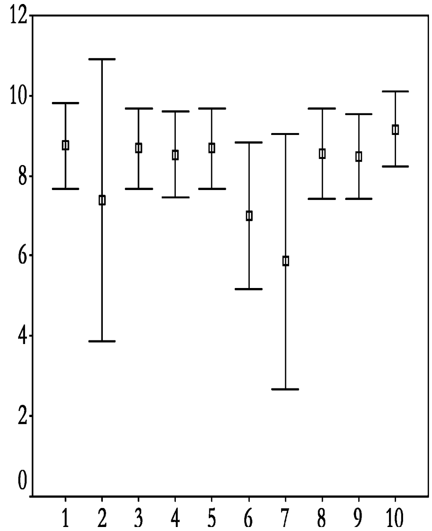

Fig. 1. Course satisfaction 95% confidence intervals.

outcome parameters

The normal distribution of the primary outcome parameter, "judgment," and the secondary outcome parameters, "fluency" and "carefulness," was confirmed by Q–Q plots.

Agreement on performance outcome

Cohen's kappa could not be computed using the original outcome scale, because it was not dichotomous or nominal, nor did the judgment outcomes fit in a similar number of classes. Compression or data into fewer classes to compute kappa as a measurement of interobserver agreement was not chosen because too much information would be lost. An alternative approach was chosen: comparing the scoring results of the two observers using nonparametrical testing (Table 3). Observers did not seem to differ in their opinion on surgeon's performance in either category.

The parameter sumscore, used to estimate the clinical outcome, was related to the final judgment. Indeed, the scores were significantly and highly correlated (Table 4).

Differences in performance outcome between experimental and control groups

Figure 3a and b graphically represent the clinical outcome as judged by both observers. Table 5 depicts the null hypothesis and its rejection for the outcome parameter "judgment" by both observers and "fluency" by observer 1.

Discussion

Published evidence on the ability to transfer skill in using VR simulation to the operating room is present,

Table 2. Satisfaction statements

1. I feel this VR-to-OR course was a valuable course for me.
2. I feel this course adequately highlights the most important aspects of the laparoscopic cholecystectomy.
3. I believe I will feel much more secure performing my first laparoscopic cholecystectomy in the OR having followed this course.
4. I think this VR-to-OR course is a highly valuable asset in the modern surgical curriculum.
5. I believe in the use of VR simulators to train surgical residents.
6. I believe in the use of VR simulators to monitor surgical residents' progress.
7. I believe in the use of VR simulators for the selection of surgical residents.
8. I think my laparoscopic surgical skills have improved significantly after following this course.
9. I felt comfortable when I did my exercises on the VR simulator.
10. I would recommend this VR-to-OR course to my colleagues.

Table 3. Agreement on the performance outcome parameters: "fluency," carefulness, and "judgment"

Group	Movement fluency	Movement carefulness	Judgment 2 vs Judgment 1
Experimental			
Z[a]	0.447	0.000	−1.509
Asymptomatic significance (two-tailed)	0.655	1.000	0.131
Control			
Z[a]	−1.134	−0.707	−0.640
Asymptomatic significance (two-tailed)	0.257	0.480	0.522

[a] Wilcoxon signed ranks test

Table 4. Correlations between the sumscore and judgment

Group		Sumscore mean	Judgment mean
Sumscore mean	Correlation coefficient	1.000	0.711
	Significance (two-tailed)	—	0.000[a]
Judgement mean	Correlation coefficient	0.711	1.000
	Significance (two-tailed)	0.000[a]	

[a] Kendall's tau_b. Correlation is significant at the 0.01 level (two-tailed asymptomotic significance)

but it is limited. Few studies have focused on the transfer of such skill [4–6, 14]. Grantcharov's [4] study randomized between surgical trainees receiving basic psychomotor VR training and a control group of surgical trainees who received no form of training whatsoever. In this study, the mean number of laparoscopic cholecystectomies performed in the experimental group was higher than in the control group. Considering the study design and the nonrandom bias introduced in this study, it is not surprising that the outcome results were positive

The study of Seymour et al. [14] randomized between basic psychomotor VR training and standard programmatic training. Unfortunately, perhaps the most critical phase in decision making in laparoscopic cholecystectomy (i.e., the clipping and cutting of the cystic duct and artery) was not performed by their experimental group. It is not clear whether the OR model porcine or human, nor is there information regarding patients' variability. The study conducted by Hyltander's et al. [6] research group used a Likert rating scale for camera navigation and instrument navigation in a porcine OR model.

The study by Hamilton et al. [5] was in fact the first to show improvement in OR performance with humans through MIST-VR training, with the video-trained control group showing no improvement. Psychomotor skills improved in both the VR and the video-trained control groups.

It seems that, independently of the study design, all the aforementioned studies show a marked decrease in the time required to perform the clinical procedure and the error reduction for their experimental group, as well as a significantly enhanced economy of movement. This must be regarded as indicating predictive validity for the transfer of psychomotor laparoscopic skill.

This study investigated the OR results for novice surgical residents after their participation in the Eindhoven virtual reality laparoscopic cholecystectomy training course. This tailored course integrates both basic and procedural VR skills training tasks to provide a multimodality approach, which is a novelty in procedural teaching. Surgical novice residents were compared with control subjects who had none or some (animal) training, but were considered able to start with the procedure by their superiors. In fact, the constitution of our case–control group refers much to the common, unstructured clinical practice of today. As in the traditional apprenticeship model, it is the expert surgeon who decides whether a resident is ready to start operating on

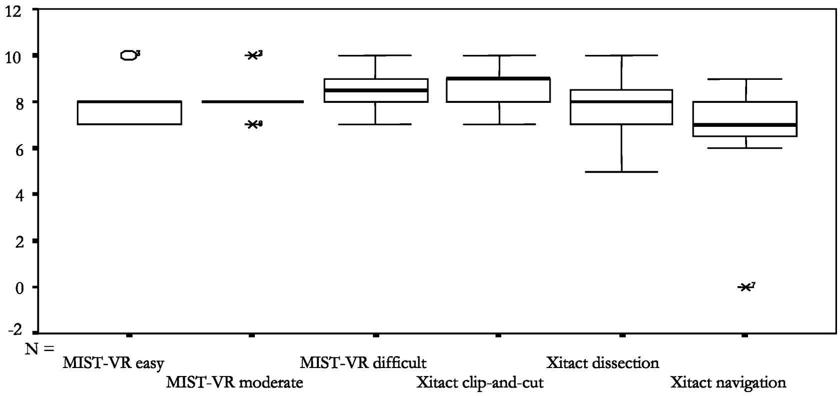

Fig. 2. **a** Clinical outcome (observer 1). **b** Clinical outcome (observer 2).

Table 5. H0: Groups do not differ in performance status

	Observer 1			Observer 2			
	Fluency	Carefulness	Judgment	Fluency	Carefulness	Judgment	Time
U[a]	24,000	37,500	12,000	29,000	31,500	18,000	33,500
Asymptomatic significance (two-tailed)	0.0037	0.214	0.004*	0.077	0.108	0.013*	0.212

[a] Mann–Whiney *U* test

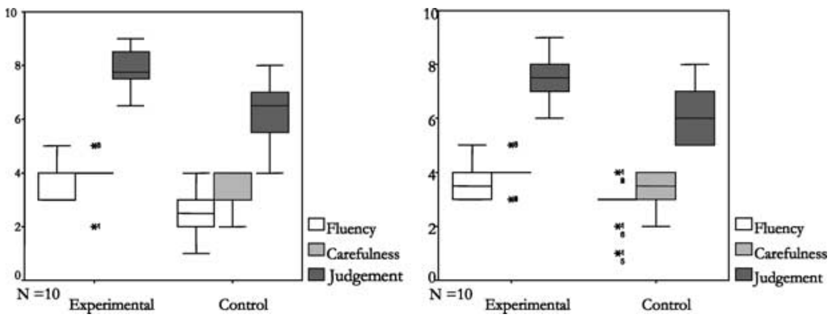

Fig. 3. Plot of "judgment" vs "time" (sec).

patients or not. The groups were demographically comparable except for the number of laparoscopic cholecystectomies performed (Table 1).

In the case–control group, there was even a somewhat more clinical experience present. However, as defined by the selection criteria, this was limited to a maximum of three clinical procedures. Therefore, the participants in both groups had to be considered well at the beginning of the learning curve [12, 16]. The fact that among the controls significantly more clinical experience was present, did not seem to contribute much to their performance outcome, as observers agreed (there were no significant differences between observers on outcome parameters) (Table 3). Rating scales were constructed for the observers because this is regarded as the most reliable and valid method for observers to use in assessing performance during a laparoscopic cholecystectomy [3].

Observers judged the overall performance in the experimental group to be clearly superior. One observer also considered the pattern of movement, in terms of fluency, to be significantly better for this group than for the performers in the control group (Table 5). The parameter "sumscore," used to estimate performance outcome in earlier validation studies of the Xitact simulation software, seemed to correlate highly with observers' judgment (Table 4). Indeed, the scores were highly correlated, so the sumscore is thought to be a reliable estimate of observers' final judgement. As for parameter "time," it cannot be said that there was a significant difference between the groups (Table 5). This is probably because of the large dispersion in outcome (i.e., large confidence interval). However, as depicted in Fig. 3, linking observers' judgment to time needed for completion of the clip-and-cut task results in less dispersion in procedural time within the experimental group, suggesting more efficient performance. As for the participants, they enjoyed and valued the course highly, according to the statements presented (Fig. 1 and Table 2). The participants, rated the clip-and-cut scenery, as developed and validated by Xitact the highest (Fig. 4) [9, 10].

A few limitations to the methods of this study must be addressed. The participants were not randomly assigned to either group. In fact, the experimental subjects were analyzed against the background of a group of controls trained according to "current clinical practice in The-Netherlands" (i.e., declared fit to perform the

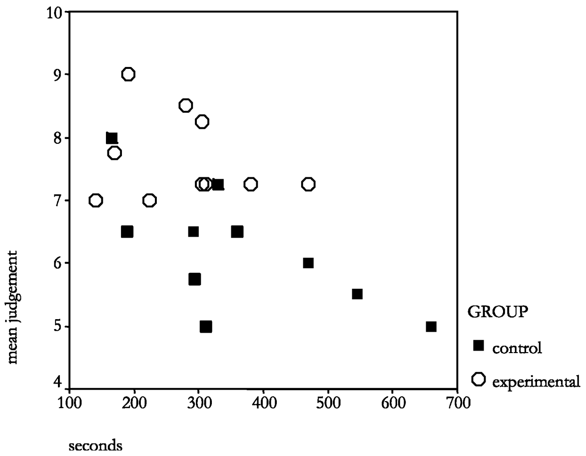

Fig. 4. Virtual reality simulation ratings (0–10 point scale).

procedure according to the opinion of their own clinical supervisor). Because there were no baseline recordings using VR simulation available for both the experimental and the control groups on the assessed clip-and-cut procedure, it cannot be firmly stated that the participants were equally skilled at inclusion. Although the inclusion criteria restricted clinical experience to a maximum of four procedures, and although the control subjects were in fact more experienced beforehand, their outcome was significantly worse. The better performance of the residents in the experimental group after the VR course must be interpreted with care because it cannot be ruled out that the control group had a lower level of skills to begin with.

Also, by assessing only the clip-and-cut part, one must be careful in extrapolating study results to the whole procedure of the laparoscopic, cholecystectomy. Nevertheless, it is highly unlikely that residents who are not fluent or careful in their motions on the clip-and-cut part are technically able to dissect Calot's triangle properly.

Nonanimal training models for acquiring endoscopic surgical skill usually are limited to box trainers. Bench models are safe, usually readily available, and inexpensive. It is believed that psychomotor skills can be trained in these environments effectively, although there is no form of objective assessment besides structured observation, and no procedural training can be performed [8, 13, 15].

Animal training models provide excellent opportunities for tissue handling and exposure to anatomy almost identical to the human counterpart. However, animal models are costly and cannot be repeated easily, thus providing inadequate possibilities for the resident to reach the plateau phase inherent to training for a new laparoscopic procedure. Also, animal training models require a demanding skills laboratory infrastructure.

Virtual reality simulation is bound to become "the next big thing" in surgery training, and more specifically in MAS. Virtual reality simulation is free of the aforementioned boundaries and limitations, provided the systems are well validated according to a consented route [10]. In fact, simulation systems have the extra asset of an objective scoring system and the inherent possibility of constructing individual learning curves [12].

Earlier studies have shown that VR training results in psychomotor skill acquisition at least as good as, if not better than, programs using conventional box trainers [7, 8]. The second generation of VR simulators provide excellent real-time graphics, various patient scenarios, and anatomic variations, combined with near-to-real haptic sensations and behavior of the simulation. Therefore, it is important for both surgeons and educators to be aware and oriented in this fast-expanding field of educational tools [11]. Precisely so, the Eindhoven course was further on developed featuring aspects of both LapSim and Simbionix LapMentor VR simulation software.

Conclusion

Procedural VR simulation can no longer be considered fancy "eye candy." Indeed, it should be regarded as a serious, cost-effective teaching instrument for acquiring endoscopic surgical skill. Our study, integrating multiple repetitive, complementary VR simulations, showed that OR performance was significantly better in the experimental group.

Therefore, the Eindhoven virtual reality laparoscopic cholecystectomy training course is likely to increase participants' level of skill such that it shortens the learning curve inherent to the procedure in a safe environment. Skills derived from the course can be brought to bear on clinical situation, and will be an asset in the curriculum used to train future surgeons.

References

1. Ahlberg G, Heikkinen T, Iselius L, Leijonmarck CE, Rutqvist J (2002) Does training in a virtual reality simulator improves surgical performance? Surg Endosc 16: 126–129
2. Bridges M, Diamond DL (1989) The financial impact of teaching surgical residents in the operating room. Am J Surg 210: 118–121
3. Eubanks TR (1999) An objective scoring system for laparoscopic cholecystectomy. J Am Coll Surg 189: 556–574
4. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J (2004) Randomized clinical trail of virtual reality simulation for laparoscopic skills training. Br J Surg 91: 146–150
5. Hamilton EC, Scott DJ, Fleming RV, Rege RV, Laycock R (2002) Comparsion of video trainer and virtual reality training systems on acquisition of laproscopic skill. Surg Endos 16: 406–411
6. 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: 1324–1328
7. Rossner JC, Rossner LE, Savalgi RS (1997) Skill acquisition and assessment for laparoscopic surgery. Arch Surg 132: 200–204
8. Rossner JC, Rossner LE, Savalgi RS (1998) Objective evaluation of laparoscopic surgical skill program for residents and senior surgeons. Arch Surg 133: 657–461
9. Schijven M, Jakimowicz J (2002) Face, expert, and referent validity of the Xitact LS500 Laparoscopy Simulator. Surg Endosc 16: 1764–1770
10. Schijven M, Jakimowicz J (2003) Construct validity: experts and residents performing on the Xitact LS500 laparoscopy simulator. Surg Endosc 17: 803–810
11. Schijven M, Jakimowicz J (2003) Virtual reality surgical laparoscopic simulators: how to choose. Surg Endsc 17: 1943–1950

12. Schijven MP, Jackimowicz JJ (2003) The learning curve on the Xitact LS500 laparoscopy simulator: profiles of performance. Surg Endosc 18: 121–127
13. Scott DJ, Young WN, Tesfay ST, Frawley WH, Rege RV (2000) Laparoscopic training on bench models: better and more cost effective than operating room experience? J Am Coll Surg 191: 272–283
14. Seymour NE, Gallagher AG, Roman SA, O'Brian MK, Bansal VK (2002) Virtual reality improves operating room performance: results of a randomized, double-blind study. Ann Surg 236: 458–463
15. Taffinder NJ, Russell RCG, MCManus IC, Jansen J, Darzi A (1998) An objective assessment of surgeons' psychomotor skills: validation of the MIST-VR laparoscopic simulator. Br J Surg 85(Supp 1): 75
16. The Southern Surgeons Club (1995) The learning curve for laparoscopic cholecystectomy. Am J Surg 170: 55–59