

Validation of a new box trainer-related tracking device: the TrEndo

Pieter J. van Empel · Lennart B. van Rijssen · Joris P. Commandeur ·
Mathilde G. E. Verdam · Judith A. Huirne · Fedde Scheele ·
H. Jaap Bonjer · W. Jeroen Meijerink

Received: 29 June 2011 / Accepted: 5 January 2012 / Published online: 21 February 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract

Background There is an increasing demand for structured objective ex vivo training and assessment of laparoscopic psychomotor skills prior to implementation of these skills in practice. The aim of this study was to establish the internal validity of the TrEndo, a motion-tracking device, for implementation on a laparoscopic box trainer.

Methods Face validity and content validity were addressed through a structured questionnaire. To assess construct validity, participants were divided into an expert group and a novice group and performed two basic laparoscopic tasks. The TrEndo recorded five motion analysis parameters (MAPs) and time.

Results Participants demonstrated a high regard for face and content validity. All recorded MAPs differed significantly between experts and novices after performing a square knot. Overall, the TrEndo correctly assigned group membership in 84.7 and 95.7% of cases based on two laparoscopic tasks.

Conclusion Face, content, and construct validities of the TrEndo were established. The TrEndo holds real potential as a (home) training device.

Keywords Box trainer · TrEndo · Laparoscopy · Training · Psychomotor skills · Objective assessment · Simulation · Motion tracking · Validity

Minimally invasive surgery (MIS) has been recognized as a contribution to the field of surgery by the majority of general surgeons, gynecologists, and urologists due to a distinct set of advantages, including lower short-term morbidity and mortality rates for laparoscopic resections compared to open surgery, cosmetic advantages, and an associated improved postoperative recovery [1, 2]. Technical–surgical demands in MIS differ from those in open surgery, including reduced depth perception [3–6], longer instruments, counterintuitive instrument movement, and loss of joint dexterity [7, 8]. The operating room (OR) as a primary, complex, and expensive teaching environment is no longer desirable and it also carries legal and ethical concerns, amplified by increasing pressure on OR efficiency [9–12]. There is a demand for structured objective ex vivo training and assessment of laparoscopic skills prior to implementation in practice.

Simulation-based practice does not put patient safety at risk and avoids interference with the efficiency of health-care resources [13, 14]. Other advantages include practice of (exclusively difficult aspects of) procedures at one's own pace and with constructive feedback. Therefore, skill acquisition is more efficient. Assessment and therefore a minimal competency level of skills prior to implementation in practice are possible. Non-patient-bound simulation-based practice has already been demonstrated to improve

P. J. van Empel (✉) · L. B. van Rijssen ·
J. P. Commandeur · J. A. Huirne · H. Jaap Bonjer ·
W. Jeroen Meijerink
Department of Surgery, VU University Medical Centre,
P.O. Box 7057, 1007 MB Amsterdam, The Netherlands
e-mail: p.vanempel@vumc.nl

M. G. E. Verdam
Department of Psychology, University of Amsterdam,
Amsterdam, The Netherlands

J. A. Huirne
Department of Obstetrics and Gynaecology, VU University
Medical Centre, Amsterdam, The Netherlands

F. Scheele
Department of Obstetrics and Gynaecology, Sint Lucas Andreas
Hospital, Amsterdam, The Netherlands

MIS skills [15, 16], which are subsequently transferable to the OR [15, 17–20]. Consequently, many surgical programs are incorporating simulation-based practice into their curriculum. However, evaluation of laparoscopic skills is currently still performed mainly by subjective expert observation [21–24].

The laparoscopic box trainer, a traditional MIS simulator, has been shown to improve MIS skills and appears to be effective as a high-fidelity training device [25–29]. The objective of this study was to investigate face, content, and construct validities of a new motion-tracking device, the TrEndo (Training in Endoscopy, Delft University of Technology, Delft, The Netherlands), implemented on a traditional laparoscopic box trainer. The TrEndo is an augmented-reality (AR) simulator that records various task-efficiency parameters (motion analysis parameters, MAPs) during simulated laparoscopic tasks. Inclusion of a motion-tracking device on the laparoscopic box trainer is a relatively new training option.

Before a surgical simulator can be used as a training and assessment device, its validity should be proven by vigorous and objective evaluation [30, 31]. *Face validity* assesses simulator realism [31–33]. *Content validity* describes the simulator's usefulness as a training tool [31, 32]. *Construct validity* reflects a simulator's abilities to discriminate between different levels of competence, e.g., experienced surgeons and novices [33–35].

Materials and methods

A prospective observational cohort study was conducted in The Netherlands and Belgium between February 1 and November 31, 2010.

Participants

Participants were divided into two groups based on prior laparoscopic experience. Experts were defined as having performed over 100 basic laparoscopic procedures and were recruited at two Dutch conferences. Medical students with no laparoscopic experience were defined as novices and were recruited at the VU University Medical Center. Trainees in urology, gynecology, and surgery participating in a laparoscopic suturing course organized by the VU University Medical Center at hospitals in The Netherlands and Belgium were additionally recruited for face and content validity evaluation purposes [36]. All participants voluntarily participated in this study. Participants with prior TrEndo experience were excluded. A brief introduction to the TrEndo was given to all participants.

Systems and hardware

Laparoscopic training boxes (Camtronics Nederland B.V., Son, The Netherlands) simulate an abdominal cavity using an aluminum frame and allow regular insertion of traditional trocars with conventional laparoscopic instruments (B. Braun Medical B.V., Melsungen, Germany) and a camera connected to a video monitor on which the simulated environment is viewed.

The TrEndo is constructed as a trocar on a laparoscopic training box through which laparoscopic instruments may still be regularly inserted (Fig. 1). Instrument movement is measured in four degrees of freedom: X, Y, and Z axes and axis rotation [37]. Five motion analysis parameters (MAPs) are recorded with a sample frequency of 100 Hz individually for the right and left hands, including: path length (mm, length of curve described by the tip of the

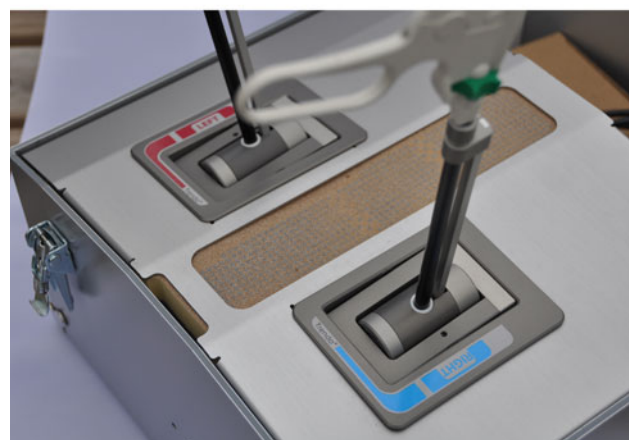
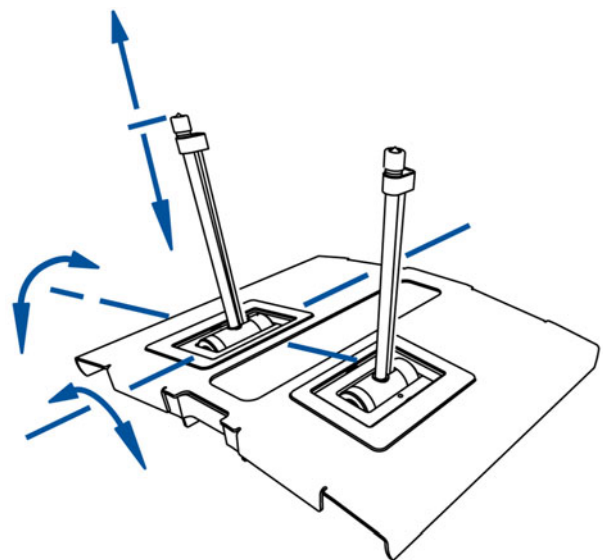


Fig. 1 **A** Schematic and photo of TrEndo motion-tracking device on a laparoscopic box trainer. **B** Knot-tying task on a laparoscopic box trainer equipped with the TrEndo tracking device



Fig. 1 continued

instrument), insertion distance (mm, total distance traveled by the instrument along its axis), angular area (deg^2 , related to the distances between the farthest positions of the instrument during a task), volume (mm^3 , three-dimensional space used), and time (s) [38].

Face and content validities

Twenty questions adapted from a previously described study on the face validation of another laparoscopic training device were used to inquire experts' and trainees' first impression of the TrEndo with a laparoscopic box trainer [33]. Participants additionally rated six possible advantages of the TrEndo with a laparoscopic box trainer compared to a virtual reality (VR) system on a 5-point Likert scale (1 = unimportant, 5 = very important) which we represent as percentages rated low–moderately important (score = 1–3) to highly important (score = 4–5). To assess content validity, participants reflected on the training capacities of the TrEndo within a standardized surgical curriculum using an open questionnaire.

Construct validity

Participants were asked to correctly position a curved tapered needle into a laparoscopic needle holder (task 1) and complete a standard laparoscopic square knot (task 2) in the laparoscopic box trainer on an artificial skin patch

using a 15-cm single 3-0 silk suture. A 5-min time limit was set for task 2. Equipment and instruments used were kept identical. Construct validity was determined by comparison of expert and novice MAPs for both tasks. To explore whether MAPs predict individual laparoscopic skill, a logistic regression analysis was conducted to predict group membership (novice or expert) based on TrEndo performance. Finally, we used a Wald analysis to determine the most contributive MAPs in predicting group membership by calculating individual contributions per MAP to the predictions made in the logistic regression analysis.

Statistical analysis

Statistical analysis was performed using SPSS 15.0 for Windows (SPSS Inc., Chicago, IL, USA). A nominal significance level of 0.05 was used. All tests were performed two-sided. Values for continuous variables are given as mean (SD), with regard to the multiple TrEndo parameters this was performed using MANOVA analysis, followed by separated ANOVAs. Comparisons with regard to the questionnaire data were performed with two-sided Mann–Whitney *U* analysis. These results are presented as median values. Values for categorical data are specified as frequency (%) and were analyzed using χ^2 or Fisher's exact test. A logistic regression analysis was used to test the predictive value of the TrEndo parameters for classification of participants as either novice or expert according to their performance.

Results

Demographics, experience, and face validity

The majority of respondents were active in general surgery (Table 1). Thirty-eight experts and 24 trainees returned the questionnaire. Not all respondents completed the entire form, but to calculate mean scores there were never more than four value points missing.

Table 2 depicts mean first impression scores. All questions on first impression were rated above a score of 3 on the 5-point Likert scale. Trainees were slightly more positive than experts, with a significant difference for design and overall functionality.

Content validity

All experts and trainees rated procedural functioning, hand–eye coordination, and depth perception above 3 on the 5-point Likert scale (Table 3). Seventy-five percent of the participants scored 4–5 on the 5-point Likert scale

Table 1 Participant characteristics

Demographics	Total	Experts	Trainees
Total (male; female)	57 (39; 18)	38 (29; 9)	24 (12; 12)
Mean age (range)		46 (33–60)	32 (28–37)
Respondent specialties	<i>N</i>	%	%
General surgery	43	71.1	66.7
Gynecology	9	13.2	16.7
Urology	10	15.8	16.7
Mean no. laparoscopic procedures/year		114.86	40.87
Mean no. complex laparoscopic procedures/year		62.97	7.05

Table 2 Face validity

	Total mean	Experts		Trainees		<i>p</i> *
		Mean	SD	Mean	SD	
Design	3.57	3.36	0.79	3.88	0.74	0.01
Realism	3.30	3.30	0.81	3.29	0.67	0.82
User-friendliness	3.69	3.68	0.66	3.71	0.69	0.81
Overall functionality	3.40	3.08	1.02	3.88	0.74	0.02
Trocar positions	3.79	3.78	0.71	3.79	0.51	0.95
Instrument movement	3.67	3.70	0.85	3.63	0.71	0.65

* Mann–Whitney test, two-sided, expert versus resident

Table 3 Content validity

Training capacities	Total mean	Experts		Trainees		<i>p</i> *
		Mean	SD	Mean	SD	
Procedural functioning	4.26	4.16	0.68	4.42	0.56	0.13
Hand–eye coordination	4.31	4.29	0.64	4.35	0.58	0.72
Depth perception	3.48	3.34	0.95	3.69	1.02	0.26

* Mann–Whitney test, two-sided, expert versus trainee

regarding didactic quality compared to a low–moderate score (1–3) by 12.5%. This question was not answered by 12.5% of participants. No significant differences were observed between experts and trainees.

Over 60% of the participants rated each selected advantage of the TrEndo with laparoscopic box trainer as highly advantageous compared to a VR system. No significant difference was observed between experts and trainees (Table 4).

Table 4 Selected possible advantages of the box trainer with TrEndo, compared to VR systems

	Total ^a		Experts ^a		Trainees ^a		<i>p</i> *
	1–3	4–5	1–3	4–5	1–3	4–5	
Intro and examples DVD	35.1	64.9	32.4	67.6	39.1	60.9	0.59
Home practice	14.0	86.0	17.6	82.4	8.7	91.3	0.91
Real instruments	5.3	94.7	5.9	94.1	4.3	95.7	0.79
Real needle and thread	0.0	100.0	0.0	100.0	0.0	100.0	1.00
Realistic haptic feedback	13.2	86.8	15.6	84.4	9.5	90.5	0.52
Objective assessment	28.1	71.9	26.5	73.5	30.4	69.6	0.74

Categories are divided into low-to-moderate (1–3) and highly (4–5) advantageous compared to a VR system. All values are expressed as percentages on a 5-point Likert scale

* χ^2 test, two-sided, expert versus trainees

^a Valid percent

Construct validity

Task 1 was completed significantly faster by experts ($n = 46$) than by novices ($n = 65$) ($p < 0.001$). Experts used significantly less path length than novices with both right and left hand and utilized a significantly smaller right-hand area and right-hand volume compared to novices. There were no significant differences between experts and novices with the remaining MAPs, although left-hand area, left-hand volume, and right-hand depth showed a trend in favor of experts.

No novices were able to complete task 2 within 5 min. Table 5 gives the MAPs for task 2, showing a significant

Table 5 Motion analysis parameters for square knot-tying task (task 2)

MAP	Novices		Experts		<i>p</i>
	Mean	SD	Mean	SD	
Left path (mm)	6,288.47	1,994.74	2,792.92	1,733.70	<0.01
Left depth (mm)	19.01	5.77	16.19	4.05	<0.01
Left area (deg ²)	142.53	86.24	89.93	55.13	<0.01
Left volume (mm ³)	2,746.13	1,790.54	1,318.09	870.48	<0.01
Right path (mm)	5780.33	2096.37	2862.95	1587.96	<0.01
Right depth (mm)	17.46	4.08	14.98	3.62	<0.01
Right area (deg ²)	104.40	51.52	85.34	40.48	0.03
Right volume (mm ³)	2,170.82	1,128.60	1,451.15	863.43	<0.01
Time (s)	277.53	39.97	117.59	77.80	<0.01

Table 6 Calculated Wald values for needle positioning (task 1) and square knot-tying task (task 2)

Dependent variable	Needle positioning		Square knot	
	Wald	<i>p</i>	Wald	<i>p</i>
Left path (mm)	10.97	<0.01	0.05	0.82
Left depth (mm)	4.63	0.03	0.09	0.76
Left area (deg ²)	0.43	0.51	2.38	0.12
Left volume (mm ³)	1.20	0.27	4.80	0.03
Right path (mm)	0.01	0.92	1.52	0.22
Right depth (mm)	0.06	0.81	0.01	0.92
Right area (deg ²)	0.27	0.61	0.38	0.54
Right volume (mm ³)	0.49	0.49	0.60	0.44
Time (s)	15.67	<0.01	15.37	<0.01

difference between novices and experts on all MAPs for both hands.

Based on recorded MAPs it was possible to correctly classify 84.8% of novices (56 of 66) and 84.4% of experts (38 of 45) for task 1. The TrEndo correctly classified 98.3% of novices (58 of 59) and 92.9% of experts (52 of 56) for task 2. Overall, the TrEndo correctly classified 84.7% of participants (94 of 111) at the first task and 95.7% of participants (110 of 115) at the second task.

Left path (Wald score 10.97), left depth (Wald score 4.63), and time (Wald score 15.67) were the most relevant parameters in determining participant level at task 1. At task 2, left volume (Wald score 4.80) and time (Wald score 15.37) were the most contributory parameters to correctly classifying participants as either expert or novice (Table 6).

Discussion

Structured objective training and assessment of surgical skills prior to implementation in the OR is a hot topic in current surgical, gynecologic, and urologic education [39]. This is one of the first studies to evaluate laparoscopic suturing tasks using a motion-tracking device on a traditional laparoscopic box trainer. Besides the laparoscopic box trainer, MIS simulators include augmented-reality (AR) and virtual reality (VR) systems [21, 32, 40, 41]. AR simulators combine the physical reality (such as in a box trainer) with virtual reality into one system. VR simulators are completely computer-based; software replicates entire MIS procedures.

AR simulators similar to the TrEndo include the Red Dragon (EDGE) (SimuLab, Seattle, WA, USA) and the ProMis (Haptica Inc., Boston, MA, USA). Both track instrument movements using a passive vision-tracking system, with cameras additionally capturing video images of internal laparoscopic instrument movement. The Red

Dragon also provides force measurement on instruments and tissue. These latter two instruments are, however, sensitive to interference and usable only in a laboratory setting [42–44]. Compared to the TrEndo they are also relatively expensive (training option) and are unable to deploy a faculty's own instruments.

Most AR and VR simulators provide objective feedback after an exercise using incorporated metrics calculated by motion tracking. Such feedback allows supervisors and trainees to monitor performance and progress objectively without supervisor's presence. Tracking systems are inherently present in VR simulators; however, their use for assessment is often not tested or validated [45]. All MIS simulators aim for a maximum realistic setting but basic laparoscopic skills training in particular remains unrealistic as AR and VR simulators are computer-based. Furthermore, most AR and VR simulators do not provide realistic tactile (haptic) feedback [46]. Previous studies illustrated the importance of haptic feedback during laparoscopic training [47, 48], demonstrating significant improvement of surgical skills in the presence of haptic feedback compared to training in the absence of haptic feedback [32, 48]. VR simulators tend to be expensive and immobile and the advanced technology compromises user-friendliness. Realism and procedural training are limited as VR simulation tends to focus on hand–eye coordination training [49].

Based on recorded MAPs, the TrEndo correctly classified 95.7% of a large and diverse study group into appropriate expert or novice groups. MAPs time, left depth, and left path were most contributory to correctly predicting group membership when positioning a needle into a needle holder, most likely because the left hand is often not used by novices during this task. After the performance of the square knot task, all MAPs differed significantly between experts and novices. Time and left-hand volume were the most contributory MAPs for this task. We believe a difference in novice versus expert dexterity is determined mainly by left-hand skill.

This study follows on the work described by Chmarra et al. [38], in which the TrEndo was able to correctly classify 23 of 31 gynecologists based on laparoscopic nonsuturing tasks, and on the work of Yamaguchi et al. [50], who demonstrated the efficacy of an objective evaluation of psychomotor skills for laparoscopic suturing using an electromagnetic motion-tracking system for MAPs time, left path length, and right speed between experienced surgeons and novice surgeons. In accordance with prior results, this study showed that time and path length help distinguish between expert and novices [22, 51]. We recruited a large and diverse study group, employed an extensive set of MAPs, and focused on laparoscopic tasks, including intracorporeal suturing and knot-tying.

Many novices had difficulty concentrating a full 5 min, and not all recruited experts had equal experience or skill in laparoscopy, and we did not acknowledge a subject's dominant hand. This may have influenced recorded MAPs.

Metrics currently must still be interpreted by a faculty member. We are investigating facilitation of this interpretation and therefore facilitate experts to novices or intermediate comparison.

Our face and content validity data indicate that the TrEndo would be acceptable to experts and trainees as a training device (including at home). Future studies should investigate the training capacities of the TrEndo. As the TrEndo is user-friendly and mobile it also holds real potential as a home training device. We are currently investigating this autonomous training option.

It is important to realize that MAPs provide an indication of task efficiency and with the exception that no novices completed a knot in this study, knot quality was not evaluated. The comparison of quantitative measures of novice to expert does provide evaluation of novice performance and does confirm construct validity. To provide an integrated objective assessment of MIS skills, efficiency parameters should be combined with other valid metrics of MIS skill. Improved efficiency parameters may correlate with improved quality parameters; however, this should be assessed in future studies. The first step in comparing quality parameters with efficiency indicators may include tasks where it is relatively easy to score task "error" performance (e.g., a stretched surgical glove where subjects must pass a needle through defined targets without tearing). Inclusion of an inexpensive tensiometer might also provide information on quality assessment; however, current tensiometers are too expensive. Besides task quality and efficiency, surgeons also need a core knowledge base, clinical decision-making and communicative skills, and the ability to think and work under stress in a team setting when performing MIS on patients [52, 53]. Future steps should assess the additional value of (recorded TrEndo) efficiency parameters into an integrated assessment of MIS skills and the objective value of (TrEndo-based) efficiency parameters herein.

Conclusion

Face, content, and construct validities of the TrEndo were established. The TrEndo holds real potential as a (home) training device.

Disclosures Pieter J. van Empel, Lennart B. van Rijssen, Joris P. Commandeur, Mathilde G. E. Verdam, Judith A. Huirne, Fedde Scheele, H. Jaap Bonjer, and W. Jeroen Meijerink have no conflicts of interest or financial ties to disclose.

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

References

1. Staudacher C, Vignali A (2010) Laparoscopic surgery for rectal cancer: the state of the art. *World J Gastrointest Surg* 2:275–282
2. Klarenbeek BR, Bergamaschi R, Veenhof AAFA, van der Peet DL, van den Broek WT, de Lange ESM, Bemelman WA, Heres P, Lacy AM, Cuesta MA (2011) Laparoscopic versus open sigmoid resection for diverticular disease: follow-up assessment of the randomized control Sigma trial. *Surg Endosc* 25:1121–1126
3. Hunter JG (1997) Advanced laparoscopic surgery. *Am J Surg* 173:14–18
4. Crothers IR, Gallagher AG, McClure N, James DT, McGuigan J (1999) Experienced laparoscopic surgeons are automated to the "fulcrum effect": an ergonomic demonstration. *Endoscopy* 31:365–369
5. Gallagher AG, McClure N, McGuigan J, Ritchie K, Sheehy NP (1998) An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. *Endoscopy* 30:617–620
6. Hanna GB, Cuschieri A (1999) Influence of the optical axis-to-target view angle on endoscopic task performance. *Surg Endosc* 13:371–375
7. Tholey G, Desai JP, Castellanos AE (2005) Force feedback plays a significant role in minimally invasive surgery: results and analysis. *Ann Surg* 241:102–109
8. Westebring-van der Putten EP, Goossens RHM, Jakimowicz JJ, Dankelman J (2008) Haptics in minimally invasive surgery—a review. *Minim Invasive Ther Allied Technol* 17:3–16
9. Bais JE, Bartelsman JF, Bonjer HJ, Cuesta MA, Go PM, Klinkenberg-Knol EC, van Lanschot JJ, Nadorp JH, Smout AJ, van der Graaf Y, Gooszen HG (2000) Laparoscopic or conventional Nissen fundoplication for gastro-oesophageal reflux disease: randomised clinical trial. *The Netherlands Antireflux Surgery Study Group. Lancet* 355:170–174
10. Gouma DJ, Go PM (1994) Bile duct injury during laparoscopic and conventional cholecystectomy. *J Am Coll Surg* 178:229–233
11. Moore MJ, Bennett CL (1995) The learning curve for laparoscopic cholecystectomy. *The Southern Surgeons Club. Am J Surg* 170:55–59
12. Sariago J, Spitzer L, Matsumoto T (1993) The "learning curve" in the performance of laparoscopic cholecystectomy. *Int Surg* 78:1–3
13. Kauvar DS, Braswell A, Brown BD, Harnisch M (2006) Influence of resident and attending surgeon seniority on operative performance in laparoscopic cholecystectomy. *J Surg Res* 132:159–163
14. Bridges M, Diamond DL (1999) The financial impact of teaching surgical residents in the operating room. *Am J Surg* 177:28–32
15. 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* 240:518–525
16. 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* 191:272–283
17. 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

18. Korndorffer JRJ, 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* 201:23–29
19. 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* 236:458–463
20. Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A (2007) Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg* 246:771–779
21. Botden SM, Jakimowicz JJ (2009) What is going on in augmented reality simulation in laparoscopic surgery? *Surg Endosc* 23:1693–1700
22. Darzi A, Smith S, Taffinder N (1999) Assessing operative skill. Needs to become more objective. *BMJ* 318:887–888
23. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 84:273–278
24. Moorthy K, Munz Y, Sarker SK, Darzi A (2003) Objective assessment of technical skills in surgery. *BMJ* 327:1032–1037
25. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, Hutchison C (1999) Assessment of technical skills transfer from the bench training model to the human model. *Am J Surg* 177:167–170
26. Grober ED, Hamstra SJ, Wanzel KR, Reznick RK, Matsumoto ED, Sidhu RS, Jarvi KA (2004) The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures. *Ann Surg* 240:374–381
27. Dhariwal AK, Prabhu RY, Dalvi AN, Supe AN (2007) Effectiveness of box trainers in laparoscopic training. *J Minim Access Surg* 3:57–63
28. Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A (2004) Laparoscopic virtual reality and box trainers: Is one superior to the other? *Surg Endosc* 18:485–494
29. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL (1998) Development of a model for training and evaluation of laparoscopic skills. *Am J Surg* 175:482–487
30. McDougall EM (2007) Validation of surgical simulators. *J Endourol* 21:244–247
31. Gallagher AG, Ritter EM, Satava RM (2003) Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc* 17:1525–1529
32. Aggarwal R, Moorthy K, Darzi A (2004) Laparoscopic skills training and assessment. *Br J Surg* 91:1549–1558
33. Schijven M, Jakimowicz J (2003) Construct validity: experts and novices performing on the Xitact LS500 laparoscopy simulator. *Surg Endosc* 17:803–810
34. Gallagher AG, Lederman AB, McGlade K, Satava RM, Smith CD (2004) Discriminative validity of the minimally invasive surgical trainer in virtual reality (MIST-VR) using criteria levels based on expert performance. *Surg Endosc* 18:660–665
35. Duffy AJ, Hogle NJ, McCarthy H, Lew JI, Egan A, Christos P, Fowler DL (2005) Construct validity for the LAPSIM laparoscopic surgical simulator. *Surg Endosc* 19:401–405
36. Advanced Suturing Course, VU University Medical Center, Amsterdam. Available at <http://www.vumc.nl/afdelingen/PAOG/Laparoscopiecourse/>. Accessed April 2007
37. Chmarra MK, Grimbergen CA, Dankelman J (2007) Systems for tracking minimally invasive surgical instruments. *Minim Invasive Ther Allied Technol* 16:328–340
38. Chmarra MK, Klein S, de Winter JCF, Jansen FW, Dankelman J (2010) Objective classification of residents based on their psychomotor laparoscopic skills. *Surg Endosc* 24:1031–1039
39. MacFadyen BVJ (2004) Teaching, training, and clinical surgery: Are we making a difference? *Surg Endosc* 18:361–362
40. Dunkin B, Adrales GL, Apelgren K, Mellinger JD (2007) Surgical simulation: a current review. *Surg Endosc* 21:357–366
41. Pearson AM, Gallagher AG, Rosser JC, Satava RM (2002) Evaluation of structured and quantitative training methods for teaching intracorporeal knot tying. *Surg Endosc* 16:130–137
42. Van Sickle KR, McClusky DA, Gallagher AG, Smith CD (2005) Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg Endosc* 19:1227–1231
43. Botden SM, Berlage JT, Schijven MP, Jakimowicz JJ (2008) Face validity study of the ProMIS augmented reality laparoscopic suturing simulator. *Surg Technol Int* 17:26–32
44. Gunther S, Rosen J, Hannaford B, Sinanan M (2007) The red DRAGON: a multi-modality system for simulation and training in minimally invasive surgery. *Stud Health Technol Inform* 125:149–154
45. Korndorffer JRJ, Kasten SJ, Downing SM (2010) A call for the utilization of consensus standards in the surgical education literature. *Am J Surg* 199:99–104
46. Basdogan C, De S, Kim J, Muniyandi M, Kim H, Srinivasan MA (2004) Haptics in minimally invasive surgical simulation and training. *IEEE Comput Graph Appl* 24:56–64
47. 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
48. Botden SM, Torab F, Buzink SN, Jakimowicz JJ (2008) The importance of haptic feedback in laparoscopic suturing training and the additive value of virtual reality simulation. *Surg Endosc* 22:1214–1222
49. Scott DJ, Dunnington GL (2008) The new ACS/APDS skills curriculum: moving the learning curve out of the operating room. *J Gastrointest Surg* 12:213–221
50. Yamaguchi S, Yoshida D, Kenmotsu H, Yasunaga T, Konishi K, Ieiri S, Nakashima H, Tanoue K, Hashizume M (2011) Objective assessment of laparoscopic suturing skills using a motion-tracking system. *Surg Endosc* 25:771–775
51. Smith SG, Torkington J, Brown TJ, Taffinder NJ, Darzi A (2002) Motion analysis. *Surg Endosc* 16:640–645
52. Wetzel CM, Kneebone RL, Woloshynowych M, Nestel D, Moorthy K, Kidd J, Darzi A (2006) The effects of stress on surgical performance. *Am J Surg* 191:5–10
53. Gallagher AG, Satava RM (2002) Virtual reality as a metric for the assessment of laparoscopic psychomotor skills. Learning curves and reliability measures. *Surg Endosc* 16:1746–1752