Surg Endosc (2002) 16: 126–129
DOI: 10.1007/s00464-001-9025-6

SURGICAL ENDOSCOPY
and Other Interventional Techniques

Does training in a virtual reality simulator improve surgical performance?

G. Ahlberg,[1] T. Heikkinen,[1] L. Iselius,[1] C.-E Leijonmarck,[2] J. Rutqvist,[3] D. Arvidsson[1]

[1] Department of Surgery, Karolinska Hospital, SE-171 76 Stockholm, Sweden
[2] Department of Surgery, St. Goran's Hospital, S-112, 81 Stockholm, Sweden
[3] Department of Pediatric Surgery, Astrid Lindgrens Children's Hospital, S-171, 76 Stockholm, Sweden

Received: 2 May 2001/Accepted in final form: 6 June 2001/Online publication: 12 November 2001

Abstract
Background: The development of computerized surgical simulators in a virtual reality environment demands models for proper validation. Recent investigations have shown that a virtual reality simulator (MIST-VR) is a reliable tool for the assessment of laparoscopic psychomotor skills and that it improves the automation of the so-called fulcrum effect. Therefore, we set out to determine whether training with the MIST-VR would improve the surgical performance of surgically inexperienced medical students and to see if results obtained in the simulator would correlate with surgical performance.
Methods: A total of 29 medical students were randomized into two groups. One group received preoperative MIST-VR training. Both groups then performed a simulated laparoscopic appendectomy in a pig. The operations were videotaped and examined by three independent observers.
Results: There was no significant difference in performance between the two groups. The performance with the MIST-VR correlated with the results in surgery.
Conclusion: A method that can measure surgical skill, based on the scoring of independent observers who view videotaped performances, seems to be reliable. MIST-VR did not improve the surgical skills of the subjects, but the results with MIST-VR did predict surgical outcome.

Key words: Surgical training — Virtual reality — Surgical simulation — MIST-VR

The demand for fewer complications in surgical procedures, together with the introduction of new technologies, such as laparoscopy, has fueled the discussion over the future of surgical education [4, 12, 15]. Moreover, in recent years, the tradition of training surgical students in animal models and human cadavers come under fire for ethical reasons [5].

The psychomotor and perceptual skills required for laparoscopic procedures differ from those needed for conventional surgery. The endoscopist has to find a way to guide his or her maneuvers in a three-dimensional environment by watching a two-dimensional screen. In other words, the operator loses the depth cue of binocularity and has to compensate with other depth cues. Another fundamental difference between laparoscopic and conventional surgery is the so-called fulcrum effect of the body wall on instrument manipulation. The tip of the instrument moves in the opposite direction to the surgeon's hand, thereby creating a dissonance between visual and proprioceptive feedback [2, 3, 4].

Comparison with other professional categories—especially pilots, who have had decades of experience with simulator training—has been a recurrent issue in recent years. It is inconceivable to imagine a pilot flying hundreds of passengers without prior practice and assessment in a flight simulator [6].

Models for surgical training and education outside the operating room are increasing, and the first generations of virtual reality simulators in the medical field are now available on the commercial market [5]. In this investigation, we studied the MIST-VR (Minimally Invasive Surgical Trainer in Virtual Reality) constructed by Virtual Presence (London, England). MIST-VR simulates, in an abstract way, six different parts of a laparoscopic cholecystectomy. It has already been shown to be a reliable assessment tool that improves the automation or the fulcrum effect [1, 2, 3]. However, one question that remains to be answered is whether training in simulators improves performance in the operating room. To answer that question, standards defining proper surgical techniques are needed. In addition, a reliable and reproducible model for validation is required [10].

The aims of this study were (a) to determine whether training with the MIST-VR simulator would improve the surgical performance of surgically inexperienced medical

Correspondence to: G. Ahlberg

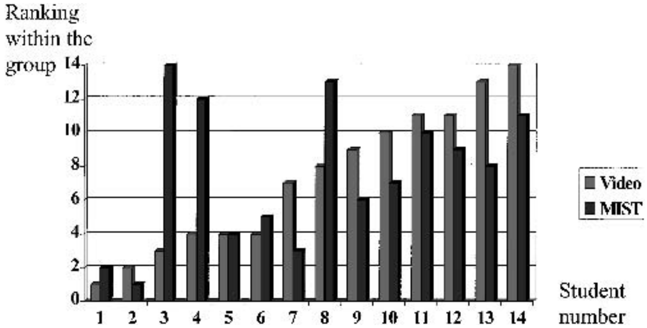

Fig. 1. Comparison of the total video scores for trained (group A) and untrained (group B) students grouped in sequences according to their scores.

students, and (b) to see whether the results obtained with the simulator would correlate with surgical performance as measured by a score calculated from a video assessment.

Methods

MIST-VR is a virtual reality simulator in a computerized environment [13, 17]. The MIST-VR system is comprised of a standard Pentium PC linked to a jig holding two laparoscopic instruments. Six different tasks simulate, in an abstract way, the maneuvers performed during a laparoscopic cholecystectomy. For each different maneuver, time measurements, errors, and economy in the movements of both hands are recorded. The system also gives proximal performance feedback to the trainee after each completed trial. Feedback is accomplished by giving each student the numerical results and relating their results to the ideal task performance, as well as noting any individual improvement.

Twenty-nine 4th-year medical students with no previous experience in endoscopy were randomized into two groups. All students volunteered and all were highly motivated. Group A (14 students) was trained under supervision with MIST-VR on two occasions preoperatively for a total of 3 h. Performance feedback was given to the students after the completion of each task. Group B (15 students) received no preoperative training.

The surgical task for each student was to perform a simulated appendectomy in a pig under the supervision of an experienced laparoscopic surgeon. A simulated appendectomy was chosen for practical reasons, since a cholecystectomy can be performed only once in each pig. All students received the same preoperative information and instruction on how the procedure should be performed. A demonstration of the operation was given by the supervisor prior to the student's own procedure.

A 3–5-cm segment of the divided small bowel was to be devascularized, looped, and divided. All operations were videotaped and examined by three independent observers. The observers did not know the identity of the student on the videotape. Five consecutive parts of each operation were judged from the videotape (grasping end of bowel, electrocoagulation of mesenteric vessels, loop ligation of bowel, cutting ligature, and division of bowel). Each part was scored as bad = 0, average = 1, or good = 2. Thus, a student could receive a score of 0–10 (maximum, 5 × 2 points) from each observer, for a maximum of 30 points (3 × 10 points).

A nonparametric test (Wilcoxon rank sum test) was used for comparison between groups. Analysis of variance (Kruskal-Wallis) was used for comparison among the three observers. A regression analysis was used to compare the video scores and MIST-VR scores of the trained group.

Results

The video scores for the two groups were compared, both as the total score (Fig. 1) for each student as well as the score for each part of the procedure (Table 1). The total score for group A was 12.0 ± 4.5 (mean ± (standard deviation [SD])); for group B, it was 12.3 ± 5.6. There was no significant difference between the groups, either for the total score or any part of the procedure.

Table 1. Comparison between the mean scores (± SD) for trained (group A) and untrained (group B) students for each part of the procedure

Maneuver	Group A	Group B
Grasping bowel	2.5 ± 1.3	2.7 ± 1.7
Coagulation of vessels	1.6 ± 1.0	2.3 ± 1.5
Loop ligation	2.1 ± 1.7	1.7 ± 1.4
Cutting ligature	3.1 ± 1.8	3.0 ± 1.4
Dividing bowel	2.6 ± 1.3	2.6 ± 1.4

For group A, video scoring was compared with the results for the most complex task in the simulator (number six version MIST-VR 1.2, where both hands and one foot have to synchronize) (Fig. 2). Except for two subjects (students 3 and 4), the results of the MIST-VR test showed a high correlation with the surgical performance (Fig. 3). Significantly, the two students who performed best in the simulator also received the highest scores in the operating room, indicating that good results in the simulator predict future potential surgical success.

Results from the three independent observers were calculated and compared (Table 2). There was no significant difference in judgment among the observers when the total score for each student was compared ($p = 0.36$). The differences seen were mainly due to variations within the group. All observers clearly picked out the same students with the best or worst performance.

Discussion

The benefits of using simulation technology in surgical education are potentially numerous. Performance can be measured and immediate feedback given without the need to have an experienced teacher present during each step of the instruction. Simulators can be made available to the students at any convenient time. The use of simulators would also minimize the expensive and problematical use of animals for educational purposes. [5, 6, 9, 11].

The field of surgical simulation is expanding rapidly. However, when validating these simulators, one must assess whether the simulator improves surgical skill in the operation theater. In this study, we could not detect any improvement in surgical skill, as measured by video scoring, after 3 h of supervised MIST-VR training.

Of course, one could argue that the training time was too short. On the other hand, the training period was approximately equal to the setup in Gallagher et al.'s series, in which 16 subjects with no experience in endoscopy were asked to make multiple defined incisions under laparoscopic laboratory conditions. Half of the subjects were randomized to receive initial MIST-VR training. According to the investigators, the simulator-trained group performed significantly better on the test due to better training in the fulcrum effect [3].

One might also argue that the software program used by us was constructed to practice cholecystectomy and not appendectomy. However, we found that the students in group A were actually slightly better only in the looping maneuver (Table 1), something not practiced in the MIST-VR. This

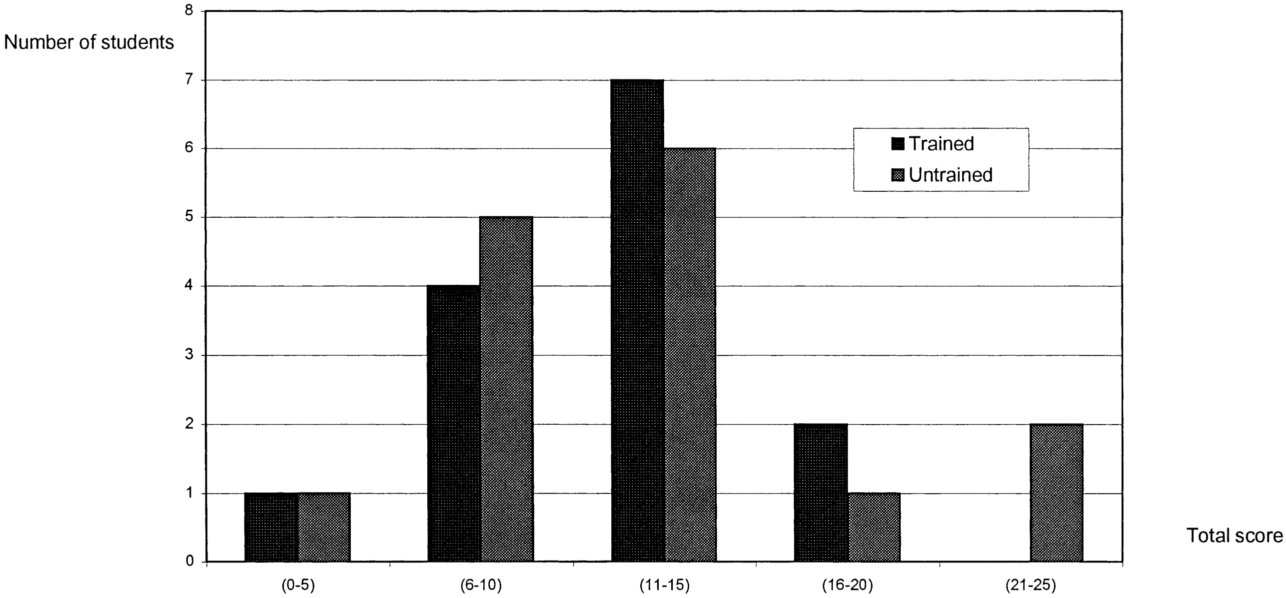

Fig. 2. Comparison of video vs MIST-VR results for the trained group (group A) given as ranking within the group. Except for two outliers (students 3 and 4), the correlation was high.

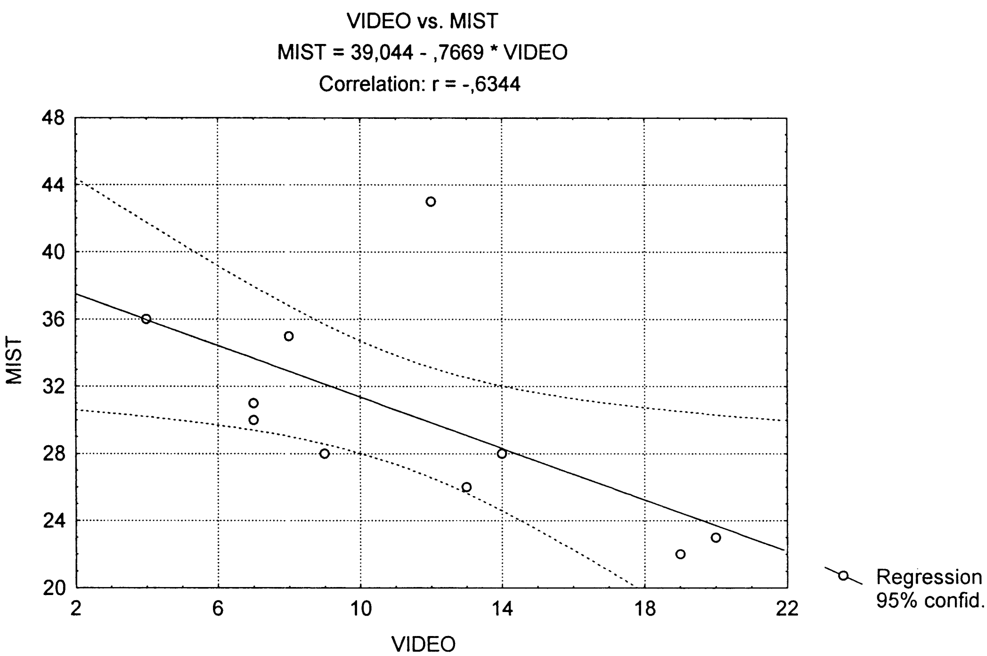

Fig. 3. Regression analysis between video and MIST-VR scores in the trained group (group A) when the two outliers (students 3 and 4) are excluded ($r = 0.64$). With outliers included, $r = 0.33$.

finding could be explained by the fact that the simulator allows the subject to practice eye-hand coordination in real time with correct dynamics but without tactile feedback in an totally abstract environment. The student is left to imagine the very complex environment involved in a laparoscopic procedure. It is likely that improvement would occur after training in simulators like MIST-VR if the evaluation were confined to tasks measuring eye–hand coordination. However, during a surgical procedure, the situation is far more complicated and depends on many other important factors, such as proper handling of the tissue, a prompt reaction to bleeding, and appropriate responses to other unexpected events [8].

Earlier studies have shown that MIST-VR is a reliable assessment tool. Taffinder et al. showed that MIST-VR clearly separated the subjects according to degree of laparoscopic experience [15, 16]. In another study, they also showed that sleep deprivation impairs cognitive performance. Wakefulness for 24 h was equivalent to a blood alcohol level of 0.10% when performance in MIST-VR was measured [14].

In this series, we found that results with MIST-VR predicted the outcome in surgery, indicating that, when used as an assessment tool, it is a reliable instrument. Interestingly enough, MIST-VR identified the best-performing pupils, who later achieved the highest scores during surgery. Whether this aspect has any predictive value for future success as a surgeon remains to be seen.

Table 2. Individual video scores for each student (1–29) as given by the independent observers (1–3)

	Student no.	Observer 1	Observer 2	Observer 3
Group A	1	3	4	6
	2	4	6	2
	3	2	0	2
	4	4	7	4
	5	3	2	3
	6	5	3	6
	7	0	5	2
	8	7	5	8
	9	2	3	4
	10	7	5	7
	11	5	4	5
	12	3	1	4
	13	4	4	6
	14	2	3	6
Group B	15	0	1	1
	16	7	7	7
	17	7	4	8
	18	5	1	2
	19	4	4	6
	20	7	8	9
	21	2	5	5
	22	5	3	2
	23	1	5	6
	24	3	4	4
	25	3	3	4
	26	1	2	3
	27	5	4	4
	28	4	3	6
	29	2	4	4

An equally important issue is the development of objective criteria for measuring surgical skill, especially when new simulators are validated [7, 16]. In this study, three independent observers, all with experience in laparoscopic surgery, evaluated the videotapes. There was no significant difference among the observers' scores; indeed, there was a high correlation among them, especially the scores for the students who performed best or worst.

In conclusion, we found that MIST-VR could predict the outcome of surgery for medical students performing a laparoscopic operation. We could not establish a positive effect of prior training in the MIST-VR, maybe due to insufficient practice time or the lack of some pedagogical aspects of the training process. We believe that simulators will become an important part of the training and assessment of surgeons in the future and that further studies in this area are necessary.

References

1. Chaundry A, Sutton C, Wood J, Stone R, McCloy R (1999) Learning rate for surgical skills on MIST-VR a virtual reality simulator: quality of human-computer interface. Ann R Coll Surg Engl 81: 281–286
2. Crothers R, Gallagher AG, McClure N, James DTD, McGuigan J (1999) Experienced laparoscopic surgeons are automated to the "Fulcrum effect": an ergonomic demonstration. Endoscopy 31: 365–369
3. Gallagher AG, McClure N, McGuigan J, Crothers I, Browning J (1999) Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST-VR). Endoscopy 31: 310–313
4. 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
5. Hoffman H, Dzung V (1997) Virtual reality: teaching tool of the twenty-first century? Acad Med 12: 1076–1081
6. Issenberg SB, McGaghie WC, Hart IR, Mayer JW, Felner JM, Petrusa ER, Waugh RA, Brown DD, Safford RR, Gessner IH, Gordon DL, Gordon AE (1999) Simulation technology for health care professional skill training and assessment. JAMA 282: 861–866
7. Johnston R, Bhoyrul S, Way L, Satava R, McGovern K, Fletcher JD, Rangel S, Loftin RB (1996) Assessing a virtual reality surgical skills simulator. Stud Health Technol Inform 29: 608–617
8. Lange T, Indelicato DJ, Rosen JM (2000) Virtual reality in surgical training. Surg Oncol Clin North Am 9: 61–79
9. O'Toole RV, Player RR, Krummel TM, Blank WC, Cornelius NH, Roberts WR, Bell WJ, Raibert M (1999) Measuring and developing suturing technique with a virtual reality surgical simulator. Am Coll Surg 1: 114–127
10. Reznick R, Regehr G, Macrae H, Martin J, McCulloch W (1996) Testing technical skill via an innovative "bench station" examination. Am J Surg 172: 226–230
11. Rosenberg LB, Stredney D (1996) A haptic interface for virtual simulation of endoscopic surgery. Stud Health Technol Inform 29: 371–387
12. Rosser JC Jr, Rosser LE, Savalgi RS (1998) Objective evaluation of a laparoscopic surgical skill program for residents and senior surgeons. Arch Surg 133: 651–657
13. Sutton C, McCloy R, Middlebrook A, Chater P, Wilson M, Stone R (1997) MIST-VR: a laparoscopic surgery procedures trainer and evaluator. Stud Health Technol Inform 39: 598–607
14. Taffinder NJ, McManus IC, Gul Y, Russel RCG, Darzi A (1998) Effect of sleep deprivation on surgeons dexterity on laparoscopic simulator. Lancet 10: 1191
15. Taffinder N, Russel 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: 75
16. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Carzi A (1998) Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST-VR laparoscopic simulator. Stud Health Technol Inform 50: 124–130
17. Wilson MS, Middlebrook A, Sutton C, Stone R, McCloy RF (1997) MIST-VR: A virtual reality trainer for laparoscopic surgery assesses performance. Ann R Coll Surg Engl 79: 403–404