Surg Endosc (2006) 20: 511–518 DOI: 10.1007/s00464-005-0230-6

- Springer Science+Business Media, Inc. 2006

and Other Interventional Techniques

Validation of a new basic virtual reality simulator for training of basic endoscopic skills

The SIMENDO

E. G. G. Verdaasdonk, ^{1,2} L. P. S. Stassen, ^{1,2} L. J. Monteny, ¹ J. Dankelman¹

¹ Man Machine Systems Group, Faculty of Mechanical, Maritime & Materials Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

 2^{2} Department of Surgery, Reinier de Graaf Group, Reiniere de Graatweg 3-11, 2625 AD Delft, The Netherlands

Received: 13 April 2005/Accepted: 7 October 2005/Online publication: 25 January 2006

Abstract

Background: The aim of this study was to establish content, face, concurrent, and the first step of construct validity of a new simulator, the SIMENDO, in order to determine its usefulness for training basic endoscopic skills.

Methods: The validation started with an explanation of the goals, content, and features of the simulator (content validity). Then, participants from eight different medical centers consisting of experts $(\geq 100$ laparoscopic procedures performed) and surgical trainees $(<100$) were informed of the goals and received a ''hands-on tour'' of the virtual reality (VR) trainer. Subsequently, they were asked to answer 28 structured questions about the simulator (face validity). Ratings were scored on a scale from 1 (very bad/useless) to 5 (excellent/very useful). Additional comments could be given as well. Furthermore, two experiments were conducted. In experiment 1, aimed at establishing concurrent validity, the training effect of a single-handed hand–eye coordination task in the simulator was compared with a similar task in a conventional box trainer and with the performance of a control group that received no training. In experiment 2 (first step of construct validity), the total score of task time, collisions, and path length of three consecutive runs in the simulator was compared between experts (>100 endoscopic procedures) and novices (no experience).

Results: A total of 75 participants (36 expert surgeons and 39 surgical trainees) filled out the questionnaire. Usefulness of tasks, features, and movement realism were scored between a mean value of 3.3 for depth perception and 4.3 for appreciation of training with the instrument. There were no significant differences be-

tween the mean values of the scores given by the experts and surgical trainees. In response to statements, 81% considered this VR trainer generally useful for training endoscopic techniques to residents, and 83% agreed that the simulator was useful to train hand–eye coordination. In experiment 1, the training effect for the single-handed task showed no significant difference between the conventional trainer and the VR simulator (concurrent validity). In experiment 2, experts scored significantly better than novices on all parameters used (construct validity).

Conclusion: Content, face, and concurrent validity of the SIMENDO is established. The simulator is considered useful for training eye–hand coordination for endoscopic surgery. The evaluated task could discriminate between the skills of experienced surgeons and novices, giving the first indication of construct validity.

Key words: Imaging $&$ VR \le Technical — Training \le Endoscopy — Education

Since the introduction of laparoscopic cholecystectomy, opinions in medical society about training of minimally invasive surgical skills have changed. There is consensus that surgical training should be structured and assessment of skills should be introduced to ensure safe and high-quality treatment [1, 8, 9, 32]. Training in the operation room (OR) is time-consuming [4, 5], and exposing the patient to relatively inexperienced surgical residents is potentially unsafe. Furthermore, recently reduced working hours for residents in The Netherlands and other countries have reduced the time available for practical training of procedures in the OR. Training surgeons according to the apprenticeship model only is Correspondence to: E. G. G. Verdaasdonk no longer acceptable [20, 22]. Therefore, most teaching

Fig.1. Left: the SIMENDO connected to the PC via USBport. Right: examples of exercises. Piling up of cylinders and clipping of a vessel.

hospitals have adopted training courses prior to training in the OR. During these courses, surgical residents train with box trainers, virtual reality (VR) trainers, or animal models. VR simulators can provide a challenging, safe, and controlled environment to master the basic skills needed to perform laparoscopic surgery [18, 23, 24, 30]. Other advantages are objective automatic scoring of performance and the possibility of unlimited repetitions of training situations.

Several VR simulators for training of laparoscopic techniques and procedures have been developed [21, 27]. Surgeons receiving VR simulator training show significantly improved performance in the OR compared to those in control groups, measured in task time and errors [16]. However, the use of VR simulators in training hospitals is limited. This may be partly due to their high cost, the extensive system requirements, and their relatively immobile characteristics. In this context, there is an increasing interest in effective, mobile, basic, and thereby affordable VR training tools for endoscopic techniques outside the OR. The SIMENDO (Delltatech, Delft, The Netherlands) is a new VR trainer developed to specifically meet these demands. The training tasks in the simulator are based on thorough assessment and research of hand–eye coordination during laparoscopic surgery [33].

Prior to implementation of a new training tool in a curriculum, evaluation and validation of the tool and its parameters are mandatory. Subjective approaches to validation include content and face validity. Content validity is generally defined as ''an estimate of the validity of a testing instrument based on a description of the contents of the test items'' or a judgment about what domains the instrument trains (e.g., psychomotor skills or anatomy) [2, 10]. Therefore, content validation is more a summation of contents of the device under study than an actual study. Face validity refers to whether the model resembles the task it is based on and addresses the questions to what extent the instrument simulates what it is supposed to represent and whether it is considered useful for training [2, 9, 10, 25]. Most studies compare the opinions of experts with those of nonexperts. In

concurrent validity, the relationship between the test scores on the trainer under evaluation and the scores achieved on another instrument purporting to measure the same construct are compared [10]. Construct validity can be defined as ''evaluating a testing instrument based on the degree to which the test items identify the quality, ability, or trait it was designed to measure'' [10]. This is usually done by measuring performance in two groups that are hypothesized to differ in the skill being measured by the instrument (e.g., experienced surgeons and novices) [7, 13, 26].

The aim of the current study was to establish content, face, and concurrent validity and perform the first step of construct validity of the SIMENDO, thereby determining its usefulness for training basic endoscopic skills.

Materials and methods

The system: hardware and system requirements

The SIMENDO (simulator for endoscopy) consists of one instrument handle on a box weighing 1.0 kg and measuring $10 \times 10 \times 40$ cm (Fig. 1). The software is integrated in the system and provides ''plugand-play'' connectability via a USB port. Users of this simulator do not need to install additional software to be able to practice with the instrument. Each PC with a Microsoft Windows XP operating system is directly accessible for the simulator. Minimal computer requirements are a 722-mHz processor, 128 MB RAM, a standard graphical card (NVIDIA Geforce 4), and Microsoft Office software (with Access database).

Content validation

Program and tasks (SIMENDO version 1.0.0)

The exercises in the training program are designed to train hand–eye coordination using abstract tasks without force feedback. The training program in the simulator starts with a short theoretical explanation of the difficulties a surgeon faces during endoscopic procedures. The goal is to train nonexpert subjects the skills needed to deal with specific characteristics of endoscopic surgery, such as the fulcrum effect, the use of long instruments, hampered depth perception, scaling of instru-

Fig.2. Distribution of participants as function of number of endoscopic procedures performed.

ments, and misorientation. During the explanation, the user is asked to manipulate a virtual endoscope and instruments to demonstrate misorientation during laparoscopic surgery [33]. Then, the user can choose between four different tasks: piling up of cylinders (Fig. 1), manipulation of a 30° endoscope, clipping an artery, and dissecting a gallbladder. Also, a game is available called ''catch the needles'' in which the skills described are practiced. All the tasks, except the 30° endoscope, can be performed at three different levels. In each level, the angle between the endoscope and the instrument is increased (augmented misorientation). Besides these levels, the user can change the distance and the angle between the instrument and endoscope in any direction.

Time needed to complete the task and the number of errors are automatically measured and displayed. The errors are predefined as collisions with nontarget structures and the inappropriate placing of a clip or the dropping of it. It is also possible to track movements and measure the path length of the instruments. The user can alter the settings for each task, such as the entry positions of instruments and the camera.

Face validation

Participants

Expert surgeons and surgical trainees from eight different hospitals in The Netherlands were introduced to the simulator between October 1 and November 30, 2004. In this study, an ''expert surgeon'' was defined as having performed ‡100 endoscopic procedures and a ''surgical trainee'' as having performed <100. The introduction to the simulator consisted of an explanation of the goals of the training system and a hands-on tour of all the components of the program. Subsequently, the participants were asked to give their opinion about the training system by completing a questionnaire.

Questionnaire

All participants were asked for their age, gender, position held in the hospital, and experience with endoscopic surgery in years and number of procedures. The opinions of the expert and trainee groups were

evaluated with 28 questions about the SIMENDO. The questions were adapted from a questionnaire previously used in a study on face validation of another VR trainer [25]. The first section of the questionnaire comprised five questions about the first impression, design, and user-friendliness of the simulator. The second section contained eight questions about the training capacities of the simulator. The questions in the first two sections had to be answered by rewarding a mark on an ordinal scale, ranging from 1 (very bad/useless) to 5 (excellent/very useful). In the third section, the participants were asked for their comments and suggestions to improve the simulator in three openended questions. In addition, one question was posed about the price of the simulator and two questions about the willingness to train with the system. The final section presented nine general statements about the suitability to train surgical residents with the simulator. These statements had to be answered with "agree," "disagree," or "no opinion.'' Participants could give additional comments for all the questions.

Concurrent validity

To establish concurrent validity, an experiment was conducted to compare performance results after training with the simulator, training with a conventional laparoscopic box trainer, and without training (control group).

Experiment 1

Twenty-four students (12 male and 12 female) with no previous experience in surgery participated. They all performed a pretest consisting of a single-handed positioning task in a box trainer. In this task, 10 points had to be touched with a laparoscopic grasper. When a nontarget surrounding was accidentally touched, a short signal sounded. Then, they were randomized into three groups of eight subjects each using sealed envelopes. The first group received training in a box trainer, the second in the VR trainer, and the third group received no training. The training in the box group consisted of dropping three cubes in holes without touching the surroundings. A

Table 1. Characteristics of participants

similar task was performed in the VR trainer (''drop the balls'' task). Both groups repeated the training task 18 times. After the training, all participants performed a posttest task that was identical to the pretest positioning task. During the pretest and posttest, the time and errors (defined as collisions with nontarget environment) were measured.

Construct validity

In order to evaluate construct validity, it was tested whether the measured parameters of a task in the VR trainer (time, collisions, and path length) could discriminate between experienced surgeons (>100 endoscopic procedures) and novices (no experience with endoscopic surgery).

Experiment 2

The first step of construct validation was performed with five experienced surgeons and 20 novices. They each performed three runs of a single-handed exercise in the VR trainer under study. The same VR task was used as in experiment 1 (drop the balls). Time, collisions, and path length were measured and saved in the database of the simulator. Data of three consecutive runs were summated for each individual and for each separate parameter. Results were compared between the experienced surgeons and novice group.

Statistical analysis

Data were analyzed using SPSS (version 11.0). Differences between the calculated mean scores of the expert and nonexpert groups were analyzed by the Kolmongorov–Smirnov test (two-sided) for the 5-point ordinal scale. Fisher's exact test (two-sided) was used to compare differences between the groups on the responses ''agree'' versus ''disagree.'' The two-tailed Mann–Whitney U-test was used to analyze differences between the nonparametric data of the groups in experiments 1 and 2.

Results

Face validity

Participants

In total, 75 surgeons and surgical trainees from eight different hospitals (three academic hospitals and five large community training hospitals) participated in this study. The ''expert'' (36/75) and ''surgical trainee'' (39/ 75) groups consisted of medical specialists and residents from the departments of surgery, gynecology, urology, and orthopedic surgery. Table 1 shows the characteristics of the participants. The majority of participants, 72% (54/75), worked in general surgery. Figure 3 shows the distribution of participants as a function of the number of endoscopic procedures performed.

First impression

Table 2 shows the mean values of the scores for the first impression of the simulator. Most values tend to be good (4), except for the correlation between the movements of the hand and the screen. The highest mean score of 4.3 was given for the appreciation of training with the device. No significant differences were found between the expert surgeons and the surgical trainee group.

Training capacities and tasks

The training capacity of endoscopic procedures in general and most of the tasks were rated good, with a mean score of approximately 4. The highest score in the category training capacities was given to training of hand– eye coordination (4.3). Training of depth perception received a relatively low score (3.3). The task ''dissection of the gallbladder'' was not considered specifically useful, as indicated by a mean score of 3.2. Table 3 provides the results of the statements. In response to the statements, 81% considered the SIMENDO useful for training of endoscopic techniques to residents in general, and 83% agreed that the simulator was useful to train hand–eye coordination. Of all participants, 91% believed that it was useful for training within the hospital, and 77% also believed that the simulator was useful for training at home. Most expert surgeons (75%) indicated that the simulator could be useful for measuring skills for endoscopic procedures. Only 40% of the trainee group agreed with this statement.

Other comments

In response to the open questions, 30 of the participants indicated a preference for a two- or three-handed simulator to train with, 23 advised on additional tasks, 14 respondents suggested including a suturing/knotting task, and nine wanted a tactile feedback added to the

Table 2. Results ratings (1 = very bad/useless and 5 = excellent/very useful)

^a Komongorov-Smirnov test, two sided, Expert vs Trainee group

Table 3. Results statements

Fisher Exact Test (2-sided) for agree vs disagree for responses from Expert vs Trainees

system. In response to the question regarding what aspects were especially liked or disliked, 13 participants stated that they liked the simplicity of the system. Eleven participants made a comment about poor depth perception, and eight disliked the fact that there was no force feedback in the device. Seventy-five percent of the

Table 4. Results other questions

Fisher Exact Test (2-sided) for yes vs no, and to high vs reasonable for responses from Experts vs Trainee

Fig.3. Mean improvement in time after training: SIMENDO (33%) vs Control (15%) $p = 0.021$ Box (42%) vs Control (15%) $p = 0.001$.

Fig.4. Mean improvement in collisions after training: SIMENDO (64%) vs Control (48%) Not Significant Box (55%) vs Control (48%) Not Significant.

participating surgeons responded that the current price of the simulator was reasonable and that they would like to have the device in their hospital (Table 4).

Concurrent validity

There were no significant differences between the performance scores of the three groups on the pretest task. Figures 3 and 4 show the results of the pretest and posttest tasks. After training, time to complete the task improved in the VR group by 33% and in the box group by 42%, which were both significantly higher compared to that of the nontraining group (15%) (Mann–Whitney U-test; $p = 0.021$ and $p = 0.001$, respectively). The number of collisions decreased in both trained groups (VR and box), but this difference was not significant compared to the nontrained group (Table 4).

Construct validity

Figure 5 gives the results for time, collisions, and path length. The boxes show the median scores of the parameters over three consecutive runs of one exercise for each group. The performance of experts was significantly better than that of the novices on all parameters. The task time was shorter [median, 102.7 sec (range, 46.7–126.7) vs 149.2 sec (range, 79.6–290); $p = 0.008$], the number of collisions less [median, 3 (range, 1–8) vs 8.5 (range, 1–21); $p = 0.038$, and the total path length shorter [median, 80.1 arbitary units (AU) (range, 50.2– 95.3) vs 94.0 arbitary units (AU) (range, 77.9–163.6); $p = 0.025$].

Discussion

The results of this study show that experts and surgical trainees believe that the VR trainer under study is a useful tool to train hand–eye coordination and basic endoscopic skills for inexperienced surgeons. Comparable reduction of time to complete an exercise is achieved with training on a conventional trainer (box trainer) and the VR trainer. Furthermore, experts outperform novices in the current VR trainer.

Structured training and assessment of surgical skills before entering the operation room and performing a procedure on a real patient is an important issue in surgical education [22, 23]. VR is considered a valuable training method for laparoscopic skills [6] and an assessment tool for objective evaluation of skill levels of trainees [11–13]. Previous studies have shown positive effects of VR training on psychomotor skills during real laparoscopic tasks [11, 15, 21, 28, 29, 31]. Only the extent to which this training should take place remains a point of discussion [17].

Unfortunately, VR simulators tend to be costly, which limits their usefulness. Another disadvantage is their relative immobility. According to most surgeons, the SIMENDO could also be used at home. Flexible,

Fig. 5. Box and whisker plots for simulation scores of novices $(n = 12)$ and experts $(n = 5)$ for total time, collisions and path length. Boxes represent interquartile range, bars represent medians and whiskers the range excluding outliers. Circles represent outliers and asterisks extreme outliers.

mobile training systems are especially interesting because several studies indicate that VR training is likely to be successful when the training schedule is intermittent rather than condensed into a shorter period of extensive practice [3, 14]. Such a schedule is most easily implemented when the simulator is easily accessible (e.g., in every teaching hospital or even at home). Advanced VR trainers can play an important role in condensed skills and assessment courses in large educational centers. Reinforcement of basic skills that diminish over time if not trained frequently, such as hand–eye coordination, can occur using simpler simulators. The SIMENDO is a VR simulator that is meant to be low priced and mobile and especially suitable for training basic skills. It can be used in a structured and gradual fashion over several intervals before the trainee takes part in more advanced courses. An advantage of VR trainers in general is that, in contrast to other simple simulators such as box trainers, improvement of performance during training is automatically recorded by registration of several parameters in a database without the need for direct observation by a researcher or faculty member. If necessary, a supervising surgeon can easily review the ''learning curve'' of the trainee in the simulator afterwards.

In general, the conceptual tasks received higher scores than the task that tried to resemble an anatomical structure, e.g., ''dissection of the gallbladder.'' Apparently, training by means of simplified anatomical structures in this simulator is not considered very useful. Some respondents advised implementation of force feedback and adding a suturing or knotting task. Currently, force feedback is not the focus of this training device because the role and implications of force feedback in laparoscopic surgery are not clear [19]. Furthermore, improving the realism of the simulation of anatomical structures, modeling a suturing or knotting task, or adding force feedback will increase cost considerably due to increased demands on the software. Such expansion of the software would reduce the simplicity of the system, and this in combination with the increased cost would exceed two primary goals of this simulator: to supply a simple, plug-and-play VR trainer at an affordable price. In addition, there is evidence that the training of conceptual tasks in VR already improves performance during laparoscopic cholecystectomy in the OR [16].

Although care was taken to optimize the design of this study, face validity contains weakness because it is based on opinions. In order to reduce this weakness, questions were adapted from a previously used questionnaire [25]. However, systemic errors can originate from the questionnaire: For example, the interpretation of questions can differ among subjects because of suboptimal formulations. Also, the enthusiasm of the presenters or the attractiveness of a new training system can bias the answers.

In addition to their opinion on validity as a training device, the participants were also asked whether the simulator could be a useful device for measuring skills in endoscopic procedures. Interestingly, in contrast to the expert surgeons, the nonexpert group tended to disagree with this statement or had no opinion on this item. This may be explained by the fact that trainees are not familiar with the measurement possibilities of VR devices in general or they may dislike the idea of accepting metrics for assessment of their performance.

Experiment 1 showed that the reduction of the time to complete the exercise was significantly higher in both trained groups compared to the control group; this was not the case for the number of collisions. Probably, it was easy to learn to avoid collisions with the environment in the pretest task, allowing for a low collision level of the control group in the posttest task.

Further studies are needed to determine the measuring capacity of the SIMENDO and its usefulness for assessment and training of basic endoscopic skills of surgical trainees in the surgical curriculum. Improvements of the simulator, such as the possibility of training with two or more simulated instruments and more tasks with better depth perception, are currently being implemented and evaluated.

Conclusion

This study showed that both expert and nonexpert surgeons considered the SIMENDO to be a useful VR training device for hand–eye coordination and basic endoscopic surgical skills. The learning effect for a simple hand–eye coordination task is comparable to the effect in the box trainer. Parameters of this task can discriminate between groups of experienced and inexperienced subjects in hand–eye coordination skills for endoscopic surgery.

References

- 1. Aggarwal R, Darzi A (2004) Surgical education and training in the new millennium. Surg Endosc 18: 1409–1410
- 2. Aggarwal R, Moorthy K, Darzi A (2004) Laparoscopic skills training and assessment. Br J Surg 91: 1549–1558
- 3. Ali MR, et al. (2002) Training the novice in laparoscopy. More challenge is better. Surg Endosc 16: 1732–1736
- 4. Babineau TJ, et al. (2004) The ''cost'' of operative training for surgical residents. Arch Surg 139: 366–370
- 5. Bridges M, Diamond DL (1999) The financial impact of teaching surgical residents in the operating room. Am J Surg 177: 28– 32
- 6. Champion HR, Gallagher AG (2003) Surgical simulation—a ''good idea whose time has come.'' Br J Surg 90: 767–768
- 7. Duffy AJ, et al. (2005) Construct validity for the LAPSIM laparoscopic surgical simulator. Surg Endosc 19: 401– 405
- 8. Emken JL, McDougall EM, Clayman RV (2004) Training and assessment of laparoscopic skills. J Soc Laparoendosc Surg 8: 195– 199
- 9. Feldman LS, Sherman V, Fried GM (2004) Using simulators to assess laparoscopic competence: ready for widespread use? Surgery 135: 28–42
- 10. 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
- 11. Gallagher AG, et al. (1999) Virtual reality training in laparoscopic surgery: a preliminary assessment of Minimally Invasive Surgical Trainer Virtual Reality (MIST VR). Endoscopy 31: 310– 313
- 12. Gallagher AG, et al. (2001) Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. World J Surg 25: 1478–1483
- 13. Gallagher AG, et al. (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
- 14. Gallagher AG, et al. (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 241: 364–372
- 15. Grantcharov TP, et al. (2001) Virtual reality computer simulation. Surg Endosc 15: 242–244
- 16. Grantcharov TP, et al. (2004) Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 91: 146–150
- 17. Haluck RS (2005) Computer-based surgical simulation is too expensive. Or is it? Surg Endosc 19: 159–160
- 18. Haluck RS, et al. (2001) Are surgery training programs ready for virtual reality? A survey of program directors in general surgery. J Am Coll Surg 193: 660–665
- 19. Heijnsdijk EA, et al. (2004) The influence of force feedback and visual feedback in grasping tissue laparoscopically. Surg Endosc 18: 980–985
- 20. Krummel TM (1998) Surgical simulation and virtual reality: the coming revolution. Ann Surg 228: 635–637
- 21. Lehmann KS, et al. (2005) A prospective randomized study to test the transfer of basic psychomotor skills from virtual reality to physical reality in a comparable training setting. Ann Surg 241: 442–449
- 22. Mac Fadyen BV Jr (2004) Teaching, training, and clinical surgery. Are we making a difference. Surg Endosc 18: 361–362
- 23. Marshall RL, et al. (2000) Practical training for postgraduate year 1 surgery residents. Am J Surg 179: 194–196
- 24. Pearson AM, et al. (2002) Evaluation of structured and quantitative training methods for teaching intracorporeal knot tying. Surg Endosc 16: 130–137
- 25. Schijven M, Jakimowicz J (2002) Face-, expert, and referent validity of the Xitact LS500 laparoscopy simulator. Surg Endosc 16: 1764–1770
- 26. Schijven M, Jakimowicz J (2003) Construct validity: experts and novices performing on the Xitact LS500 laparoscopy simulator. Surg Endosc 17: 803–810
- 27. Schijven M, Jakimowicz J (2003) Virtual reality surgical laparoscopic simulators. Surg Endosc 17: 1943–1950
- 28. Seymour NE, et al. (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 236: 458–464
- 29. Taffinder N, et al. (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
- 30. Torkington J, et al. (2000) The role of simulation in surgical training. Ann R Coll Surg Engl 82: 88–94
- 31. Torkington J, et al. (2001) Skill transfer from virtual reality to a real laparoscopic task. Surg Endosc 15: 1076–1079
- 32. Villegas L, et al. (2003) Laparoscopic skills training. Surg Endosc 17: 1879–1888
- 33. Wentink M (2003) Hand–eye coordination in minimally invasive surgery. Theory, surgical practice & training. Faculty of Mechanical Engineering and Marine Technology, Delft University of Technology, Delft, The Netherlands