Surg Endosc (2009) 23:2026–2033
DOI 10.1007/s00464-008-9927-7

Evaluation of a new virtual-reality training simulator for hysteroscopy

Michael Bajka · Stefan Tuchschmid · Matthias Streich · Daniel Fink · Gábor Székely · Matthias Harders

Received: 6 December 2007 / Accepted: 18 February 2008 / Published online: 24 April 2008

Abstract

Background To determine realism and training capacity of HystSim, a new virtual-reality simulator for the training of hysteroscopic interventions.

Methods Sixty-two gynaecological surgeons with various levels of expertise were interviewed at the 13th Practical Course in Gynaecologic Endoscopy in Davos, Switzerland. All participants received a 20-min hands-on training on the simulator and filled out a four-page questionnaire. Twenty-three questions with respect to the realism of the simulation and the training capacity were answered on a seven-point Likert scale along with 11 agree–disagree statements concerning the HystSim training in general.

Results Twenty-six participants had performed more than 50 hysteroscopies ("experts") and 36 equal to or fewer than 50 ("novices"). Four of 60 (6.6%) responding participants judged the overall impression as "7 – absolutely realistic", 40 (66.6%) as "6 – realistic", and 16 (26.6%) as "5 – somewhat realistic". Novices (6.48; 95% confidence interval [CI] 6.28–6.7) rated the overall training capacity significantly higher than experts (6.08; 95% CI 5.85–6.3), however, high-grade acceptance was found in both groups. In response to the statements, 95.2% believe that HystSim allows procedural training of diagnostic and therapeutic hysteroscopy, and 85.5% suggest that HystSim training should be offered to all novices before performing surgery on real patients.

Conclusion Face validity has been established for a new hysteroscopic surgery simulator. Potential trainees and trainers assess it to be a realistic and useful tool for the training of hysteroscopy. Further systematic validation studies are needed to clarify how this system can be optimally integrated into the gynaecological curriculum.

M. Bajka (✉) · M. Streich · D. Fink
Clinic of Gynaecology, Dept. OB/GYN, University Hospital Zurich, 8091 Zurich, Switzerland
e-mail: michael.bajka@hin.ch

S. Tuchschmid · G. Székely · M. Harders
Computer Vision Laboratory, ETH Zurich, Zurich, Switzerland

Keywords Virtual reality · Training · Simulation · Hysteroscopy · Evaluation

Virtual reality (VR)-based surgical training systems [1] have been reported to perform at least equally well as traditional training and assessment methods such as direct observation, animal models, videotapes, and procedure logs in terms of reliability and validity [2]. During the last 10 years, a number of commercially available VR-based surgical simulators for laparoscopy have been developed and evaluated [3]. The European Association of Endoscopic Surgeons (EAES) actively promotes the acceptance of VR simulators by accrediting training courses [4] and providing guidelines for the validation of VR-based training systems [5]. In the USA, organizations such as the Society for Simulation in Healthcare (SSIH) or Advanced Initiatives in Medical Simulation (AIMS) help to push public awareness of surgical simulation.

In contrast to laparoscopy, the teaching of hysteroscopic interventions has received only little attention, with work focusing mainly on the development of physical models and box simulators [6]. Nevertheless, a team from the University of Washington has put its focus on teaching in obstetrics and gynaecology by proposing a new curriculum for hysteroscopy [7, 8]. Therein, each resident is given the opportunity to practise the resection of a large polyp on an inanimate model. Work on PC-based surgical simulation for hysteroscopy was initiated in 1998 by the NASA/Stanford National Biocomputation Center and the Stanford University Medical Media and Information Technologies (SUMMIT) group, resulting in the only hysteroscopy system commercially available to date: the AccuTouch system (Immersion Medical, Gaithersburg, Maryland, USA).

The HystSim (Hysteroscopic Surgery Simulator System, http://www.hystsim.ethz.ch) project was initiated in 2001 by the National Center of Competence in Research Co-Me (Computer Aided and Image Guided Medical Interventions, http://www.co-me.ch) of the Swiss National Science Foundation in order to build the most realistic simulator possible for hysteroscopic interventions by today's VR technology. For the successful integration of the developed simulation system into the training curriculum, its validity has to be proven through rigorous evaluation. Even if a widely accepted evaluation cascade has not been established yet, usually *face validity*, defined as "the extent to which the examination resembles real life situations" [2], provides the first and fundamental step in simulator validation [9–12]. Thereon, *construct validity*, defined as "a set of procedures for evaluating a testing instrument based on the degree to which the test items identify the quality, ability and trait it was designed to measure" [13], has to be investigated. Typically, it is established by comparing the performance on the virtual-reality simulator for groups of surgeons with different degrees of experience. The last and most significant step of the validation cascade is *predictive validity*, defined as "the extent to which the scores on a test are predictive of actual performance" [13]. However, predictive validity has only been established for a small number of surgical simulators [14–16], mainly because of the high expense and complexity of the involved virtual reality to operation room (VR to OR) study.

As a first step towards the establishment of the role of the device in the gynaecological curriculum the question of to what extent the simulation is realistic and whether novice and expert surgeons consider it useful for training has to be answered. Thus, this study evaluates the face validity of the HystSim system. Obviously, both expert and novice clinicians must accept the simulation as a realistic and useful training aid. Expert clinicians are often teaching hysteroscopic interventions and rely on a realistic simulation in order to expose the novice surgeon to a wide range of situations, while novices need to be willing to work with the system. It is therefore important to know the overall acceptance as well as finding out whether novice and expert surgeons perceive the simulation differently.

Material and methods

In this study we investigate the acceptance of the HystSim training simulator based on the responses of potential trainers and trainees. The concept of comparing expert and novice face validity as a first step in simulator validation has been implemented by various other validation studies in endoscopy [10, 17] and laparoscopic surgery simulation [11, 12]. Questionnaires were used to judge realism and training capacity on five-, seven- or ten-point Likert scale with the aid of agree–disagree statements. For an optimal tradeoff between differentiation and observability, we decided to use a seven-point Likert scale as proposed in [17]. Formal exemption of the institutional review board had been obtained so approval for the study was not required.

Subjects

Sixty-two gynaecological surgeons with various levels of expertise and no prior exposure to the HystSim system were interviewed at the 13th Annual Practical Course in Gynaecologic Endoscopy in Davos, Switzerland, March 1–4, 2007, organized by Gynécologie Suisse (Swiss Society of Obstetrics and Gynaecology). In accordance with the logbook of the Swiss Medical Society FMH (Foederatio Medicorum Helveticorum) to complete specialization in OB/GYN, participants having performed more than 50 hysteroscopies were defined as "experts", while those having performed equal to or fewer than 50 interventions were defined as "novices". All participants received a 20-min hands-on training on the simulator including two different diagnostic interventions, rollerball ablation, and myomectomy. Before answering the questionnaire, an informed consent form was signed, stating that the data gained from the questionnaire may be used for scientific and/or teaching purposes.

Questionnaire

The questionnaire consisted of a four-page survey divided into the categories personal settings, realism, training, and statements. For the demographic questions, the basic items concerning age, gender, dexterity, and prior experience were extended with questions on prior experience with surgical simulators in general and previously attended skills training courses or other educational programs. Fourteen questions with respect to the realism of the simulation were answered on a seven-point Likert scale: "1 – absolutely not realistic", "2 – not realistic", "3 – somewhat not realistic", "4 – undecided", "5 – somewhat realistic", "6 – realistic", and "7 – absolutely realistic". Nine questions concerning the training usefulness of the simulation were rated on a seven-point Likert scale ranging from "1 – strongly disagree" to "7 – strongly agree" with similar intermediate steps. In addition, participants answered 11 agree–disagree statements concerning training with the simulator.

Apparatus

The HystSim virtual-reality system consist of four major components: a covered mannequin pelvis, an adapted resectoscope, computing hardware, and simulation software. It allows procedural training of diagnostic and therapeutic interventions. The haptic interface is integrated into the mannequin's pelvis and offers all degrees of freedom needed, notably spherical displacement around the virtual cervix as well as translation along and rotation around the tool axis [18]. The adapted resectoscope includes the valves for controlling in- and outflow of the virtual distension fluid, tracks the displacement of the operative electrode, and provides camera rotation and adjustment of focus. Thus, all typical manipulations of the original tool can be performed in the immersive setup. Figure 1 shows the HystSim setup used in this study.

Fig. 1 Hysteroscopic surgery simulation setup for hands-on training. An adapted resectoscope is used to control the virtual-reality simulation and allows the procedural training of interventions such as hysteroscopic myomectomy

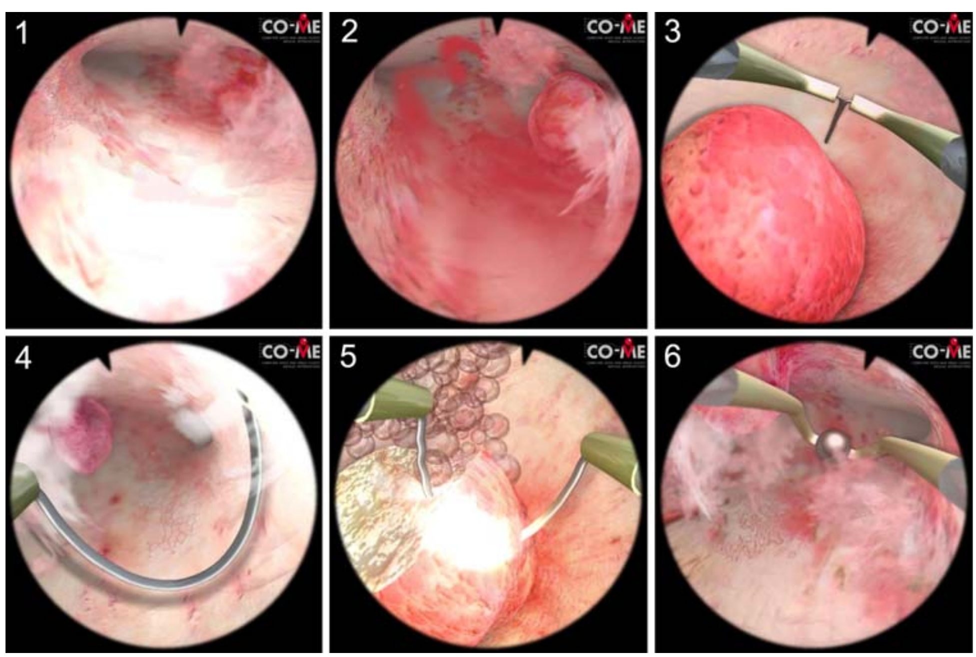
Software and hardware

Different patient cases with varying pathologies have been defined in order to replicate the day-to-day workload of a surgeon [19, 20]. Triggered by the insertion of the resectoscope, the visual display of the virtual scene starts automatically with a view obtained after passing the cervix. Figure 2 displays screenshots of different running training scenes. The simulation software runs on standard personal computer (PC) hardware (dual 3.0 GHz Pentium processor, 2 GB RAM, NVIDIA 8800 Graphics Card).

Performance measurements

Validated performance assessment on surgery simulators requires objective measures [21]. Therefore, we implemented metrics for diagnostic hysteroscopy including the quantification of the properly visualized surface as well as the quality of the endoscopic view, the safe handling of the hysteroscope, and the amount of virtual distension media used [22]. The movements of the hysteroscope are tracked, safety- and economy-related parameters of the gestures are computed, and all collisions between the instrument and

Fig. 2 Selected screenshots from sample diagnostic (1–2) and therapeutic (3–6) hysteroscopic interventions. Training scenarios 1 and 2 (diagnostics), 5 (myomectomy), and 6 (rollerball ablation) were chosen for the study



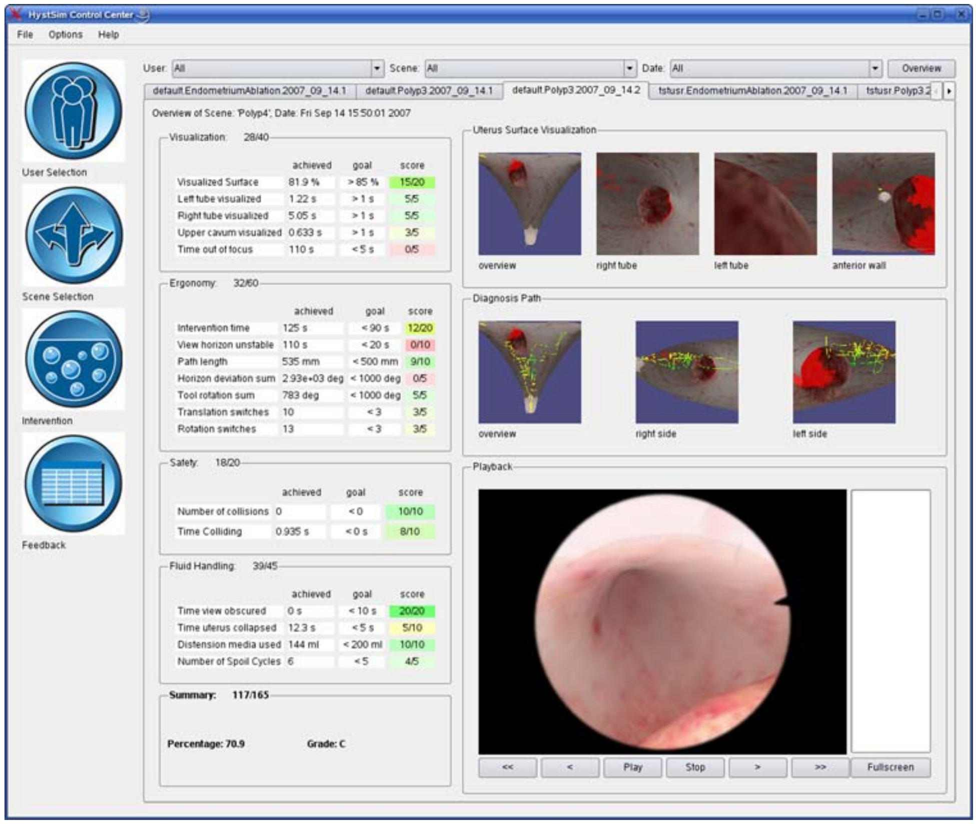


Fig. 3 Left: objective surgical performance assessment in the form of an automatically generated intervention report with various performance metrics. Upper right: uterine surface patches not visualized are marked in red. Lower right: complete video recording of a trial, which may be used for stepwise procedure analysis

the tissue are registered. The virtual fluid consumption is determined in real time by simulating the effects of leakage at the cervix, loss at the continuous flow sheath interface, outflow suction, transtubal loss, and intravasation. Finally, structured feedback is given by displaying selected criteria as postprocedure report of the intervention (Fig. 3) to stimulate improvement between repeated trials.

Statistics

A prior sample size calculation for the null hypothesis (expert opinion is equal to novice opinion) with acceptable significance (type I error, alpha = 0.05) and acceptable type II error (beta = 0.20, power = 0.8), relying on an estimate of standard deviation taken from similar studies [11, 12] (sigma = 1.0) and targeting minimum absolute difference between mean values to detect ($D = 1.0$), found a required sample size of 16 for both groups (two-sided, independent). Data were analyzed using the Statistical Package for the Social Sciences SPSS version 14.0 for Windows (SPSS Inc., Chicago IL, USA). For the questions about simulator realism and training capacity, the novice and the expert group were compared using the Mann–Whitney U-test to check for the significance of the differences. Fisher's exact test (two-sided) was used to compare outcomes for the groups on the responses "agree" versus "disagree" for the statements section. A P-value of less than 0.05 was considered as significant.

Results

Demographics

Classification of the participants resulted in 26 (42%) expert surgeons and 36 (58%) novices (Fig. 4). While all participants filled out the questionnaire, not all individual

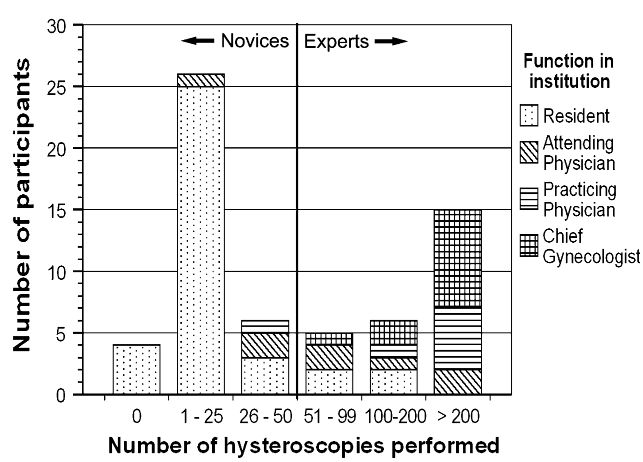


Fig. 4 Participants divided by function in institution and number of hysteroscopies performed. "Experts" were defined as having performed more than 50, "Novices" as having performed equal to or fewer than 50 interventions

questions were answered by everybody. In the following, the percentages given are calculated based on the number of respondents for the respective question. The median age of the participants was 32 years, ranging from 26 to 62 years. Forty-three (69%) were female and 19 (31%) male. Thirty-six (58%) were residents, 8 (13%) attending physicians, 11 (18%) chief physicians and heads of departments in gynaecology, and 7 (11%) practising physicians. Four of the residents had performed more than 50 hysteroscopies and were thus classified as experts, while three of the attending and one practising physician were classified as novices due to limited hysteroscopic experience. Of all participants, 52 (84%) were from German-, 8 (13%) from French-, and 2 (3%) from Italian-speaking institutions. The median experience in the field of gynaecology was 3 years (range 0–20 years). Thirty-one novices and 11 experts had not yet experienced any major complication (i.e., heavy bleeding, perforation, fluid overload syndrome). Four novices and 11 experts had mastered one to five major complications, and four experts more than five major complications. Twenty-four participants (39%) had previously attended a training course or education program in hysteroscopy, 25 (46%) in minimal invasive surgery, and 19 (35%) in any other kind of surgery. Nine (14.5%) of the participants had prior experience with surgical simulators.

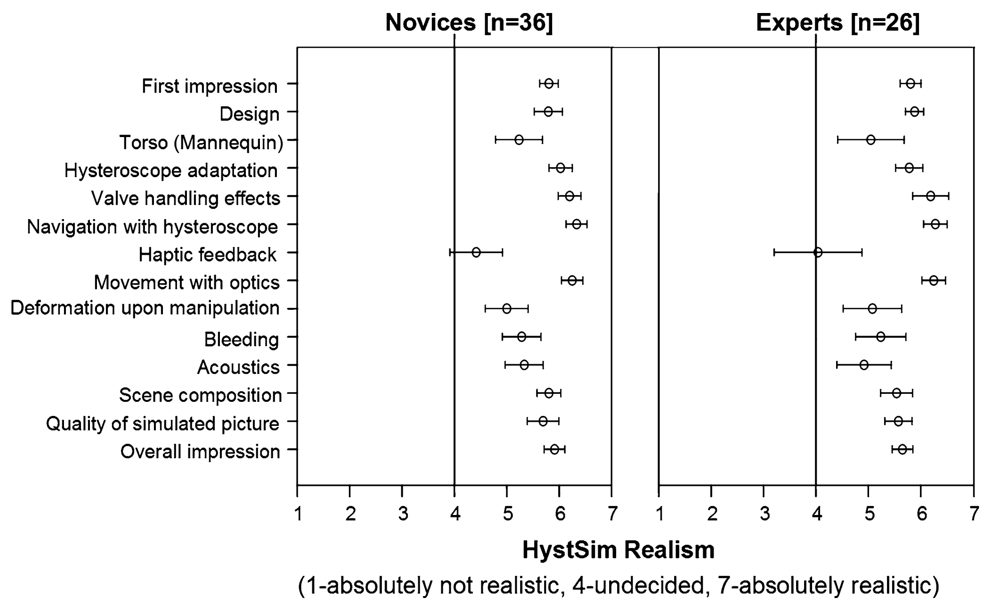


Fig. 5 Realism of the presented simulation. The mean scores for all parameters are shown with 95% confidence intervals represented by the bars. Scores are based on a seven-point Likert scale. There was no significant difference between novice and expert opinion for any aspect of the simulation ($P < 0.05$, Mann–Whitney *U*-test, two-sided, exact)

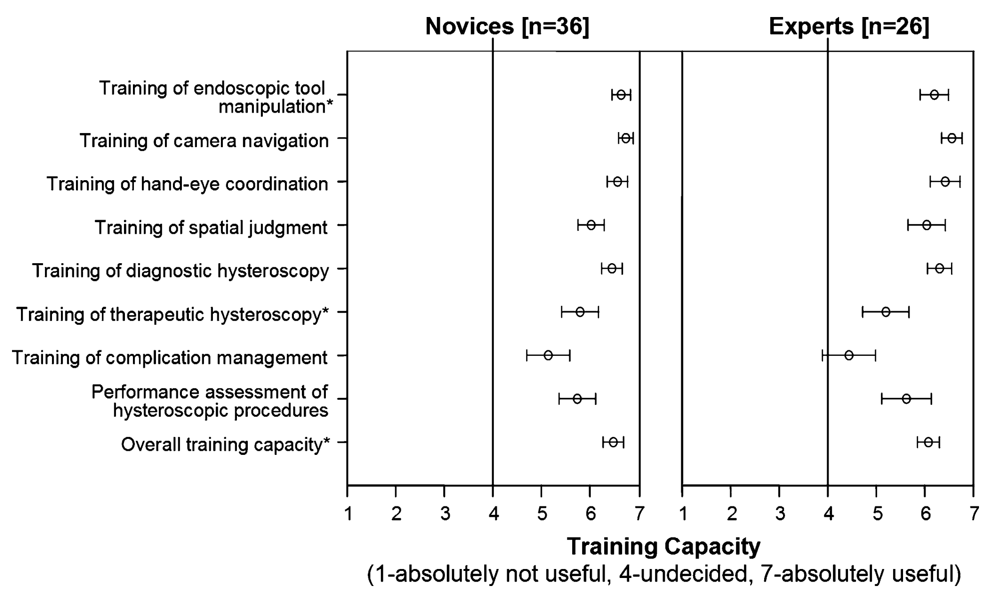


Fig. 6 Training capacity of the presented simulation. The mean scores for all parameters are shown with 95% confidence intervals represented by the bars. Scores are based on a seven-point Likert scale. * Significant difference between novice and expert opinion ($P < 0.05$, Mann–Whitney *U*-test, two-sided, exact)

Face validity

Sixteen (27%) of the participants judged the overall impression as "5 – somewhat realistic", 40 (66%) as "6 – realistic", and 4 (7%) as "7 – absolutely realistic". None of the participants ranked the simulation lower, while two participants did not answer this question. The evaluation of the different aspects of realism is displayed in Fig. 5. There was no significant difference between expert and novice opinion for all aspects of realism for the presented simulation. "Navigation with hysteroscope" (mean 6.36; 95% confidence interval [CI] 6.18–6.53) and "movement with 30° optics" (6.24; 6.07–6.42 CI) was scored highest, emphasizing the benefit of using an adapted real resectoscope. The lowest score was for "haptic feedback (tactile sensation)" which showed a high interrater variability (4.36; 3.85–4.86 CI).

Training capacity

The overall training capacity was rated as "6 – useful" or "7 – absolutely useful" by 56 of 61 participants answering this question. Novices scored the overall training capacity significantly higher (6.48; 6.28–6.7 CI) than experts (6.08;

Table 1 Results statements for novice and expert opinion

Statement	Answer	Total ($N = 62$) [%]	Novices ($n = 36$) [%]	Experts ($n = 26$) [%]	p[a]
HystSim allows procedural training of diagnostic and therapeutic hysteroscopy	Agree	95.2	94.4	96.2	[b]
	Disagree	0.0	0.0	0.0	
	No answer	4.8	5.6	3.8	
HystSim offers a helpful preparation for hysteroscopic surgery	Agree	93.5	94.4	92.3	[b]
	Disagree	0.0	0.0	0.0	
	No answer	6.5	5.6	7.7	
HystSim training should be offered to all novices for training before performing surgery on real patients	Agree	85.5	80.6	92.3	0.252
	Disagree	4.8	8.3	0.0	
	No answer	9.7	11.1	7.7	
HystSim training should be recommended for any gynaecological resident to improve his/ her skills individually	Agree	90.3	100.0	76.9	0.140
	Disagree	3.2	0.0	7.7	
	No answer	6.5	0.0	15.4	
HystSim training should be integrated into the current curriculum of the specialization program of gynaecologists	Agree	58.1	61.1	53.8	1.000
	Disagree	27.4	27.8	26.9	
	No answer	14.5	11.1	19.2	
Force feedback is an important component of HystSim and should not be omitted	Agree	64.5	72.2	53.8	0.350
	Disagree	8.1	5.6	11.5	
	No answer	27.4	22.2	34.6	
There is a need for further development in the HystSim software (scenes, etc.)	Agree	85.5	86.1	84.6	1.000
	Disagree	1.6	2.8	0.0	
	No answer	12.9	11.1	15.4	
There is a need for further development in the HystSim hardware (torso, force feedback, etc.)	Agree	50.0	55.6	42.3	0.255
	Disagree	30.6	25.0	38.5	
	No answer	19.4	19.4	19.2	
I would utilize HystSim for education and training purposes in my hospital	Agree	91.9	94.4	88.5	[b]
	Disagree	0.0	0.0	0.0	
	No answer	8.1	5.6	11.5	
I would like to have HystSim in my institution	Agree	72.6	88.9	50.0	0.011
	Disagree	6.5	0.0	15.4	
	No answer	21.0	11.1	34.6	
I would recommend HystSim to my friends	Agree	93.5	97.2	88.5	[b]
	Disagree	0.0	0.0	0.0	
	No answer	6.5	2.8	11.5	

[a] Fisher's exact test (two-sided) for agree versus disagree for responses from novices versus experts. [b] Fisher's exact test is trivial since no disagree statements were selected

5.85–6.3 CI, $P = 0.008$). Figure 6 depicts the training capacity for several aspects. Other significant differences in expert versus novice opinion were on “training of endoscopic tool manipulation” ($P = 0.006$) and “training of therapeutic hysteroscopy” ($P = 0.041$), where novices rated the training capacity higher than experts.

According to the respondents, HystSim allows procedural training of diagnostic procedures (95.2% agree, 4.8% no answer) and offers helpful preparation for hysteroscopic surgery (93.5% agree, 6.5% no answer). None of the participants disagreed with the two statements. The answers to the statement section are evaluated in Table 1.

Discussion

In this study, face validity for a hysteroscopic surgery simulator (HystSim) has been established with excellent ratings for both realism and training capacity. The presented results clearly demonstrate that potential trainees and trainers accept HystSim as a realistic tool for the training of hysteroscopic interventions. There was only a small difference between expert and novice opinion. A comparison with other face validation studies is difficult because of the lack of a standardized questionnaire. Nevertheless, the ratings for realism and training capacity are at least as convincing as presented in other studies for commercially available simulator systems [10–12]. Of course, further studies will be required to validate the assessment metrics and to quantify the training effect prior to a possible integration into the gynaecological curriculum.

Haptic feedback is a controversial topic in surgical simulation because of the unclear benefit for training effectiveness and the significant costs involved. This is also reflected in the relatively low acceptance rate found in this study, where only 14 out of 26 experts agreed that force feedback is an important component of HystSim. However, the enabled force feedback was an early version, and it remains open whether a more realistic rendering of forces would have resulted in a different acceptance among participants.

Even though the study was carefully designed, face validity is a very subjective type of validation and might be influenced by systemic and nonrandom sources of errors. Opinions might be biased by the individual attention given to the participant and favorable responses can result from an enthusiastic presentation of the system. Also, the mere novelty of the simulator and its status as research project might have influenced respondents to see the presented simulation in a more favorable light. Systematic errors might also be introduced by the questionnaire. The interpretation of questions can differ among subjects, especially since the language of the questionnaire was English while the participant’s mother tongue was German, French or Italian. Also, Likert scales may be subject to distortions, e.g., by the tendency to avoid using extreme response categories or agreeing with statements as presented. We tried to minimize the error sources by developing the questionnaire with support from a social scientist, handling the questionnaire anonymously and separately from the informed consent form, and pointing out to all participants that both positive and negative feedback will be equally helpful for the future development of the simulator.

In a commentary on patient safety [23], Dr. Pearlman stresses the importance of incorporating a culture of safety at all levels of education. Virtual-reality simulation can teach basic hysteroscopic skills in the early stage of training and procedural skills during the third and fourth year of residency. Also, it might help residents in obstetrics and gynaecology to find out whether their strengths are in minimal-invasive surgery. In another commentary by Dr. Fenner [24], the importance of adapting the current methods of teaching, learning, and maintaining surgical competency is emphasized in order to meet emerging challenges such as the increased public awareness of medical errors and the mandated decrease in resident work hours, aggravated by the rapid introduction of new surgical technologies. The encouraging results of this study demonstrate that the HystSim has the potential to be a useful tool in the proposed shift towards technology supported surgical education and objective skills assessment.

Acknowledgements The authors would like to thank all developers of the hysteroscopy simulator project. This research has been supported by the NCCR Co-Me (Computer Aided and Image Guided Medical Interventions) of the Swiss National Science Foundation (http://www.co-me.ch).

References

1. Basdogan C, Sedef M, Harders M, Wesarg S (2007) Virtual reality supported simulators for training in minimally invasive surgery. IEEE Comput Graph Appl 27:54–66
2. Moorthy K, Munz Y, Sarker S, Darzi A (2003) Objective assessment of technical skills in surgery. BMJ 327:1032–1037
3. Schijven M, Jakimowicz J (2003) Virtual reality surgical laparoscopic simulators. Surg Endosc 17:1943–1950
4. Schijven MP, Jakimowicz JJ, Broeders IAMJ, Tseng LNL (2005) The Eindhoven laparoscopic cholecystectomy training course. Surg Endosc 19:1220–1226
5. Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB (2005) Consensus guidelines for validation of virtual reality surgical simulators. Surg Endosc 19:1523–1532
6. Aydeniz B, Meyer A, Posten J, König M, Wallwiener D, Kurek R (2000) The ‘HysteroTrainer’—an in vitro simulator for hysteroscopy and falloposcopy. Contrib Gynecol Obstet 20:171–181
7. Mandel LP, Lentz GM, Goff BA (2000) Teaching and evaluating surgical skills. Obstet Gynecol 95:783–785
8. VanBlaricom AL, Goff BA, Chinn M, Icasiano MM, Nielsen P, Mandel L (2005) A new curriculum for hysteroscopy training as

demonstrated by an objective structured assessment of technical skills (OSATS). Am J Obstet Gynecol 193:1856–1865
9. Aggarwal R, Tully A, Grantcharov T, Larsen CR, Miskry T, Farthing A (2006) Virtual reality simulation training can improve technical skills during laparoscopic salpingectomy for ectopic pregnancy. BJOG 113:1382–1387
10. Verdaasdonk EGG, Stassen LPS, Monteny LJ, Dankelman J (2006) Validation of a new basic virtual reality simulator for training of basic endoscopic skills: the SIMENDO. Surg Endosc 20:511–518
11. Schijven M, Jakimowicz J (2002) Face-, expert, and referent validity of the Xitact LS500 laparoscopy simulator. Surg Endosc 16:1764–1770
12. Ayodeji I, Schijven M, Jakimowicz J, Greve J (2007) Face validation of the Simbionix LAP Mentor virtual reality training module and its applicability in the surgical curriculum. Surg Endosc 21:1641–1649
13. 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
14. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK (2002) Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 236:458–463
15. Hart R, Doherty DA, Karthigasu K, Garry R (2006) The value of virtual reality-simulator training in the development of laparoscopic surgical skills. J Minim Invasive Gynecol 13:126–133
16. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Jensen PF (2004) Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 91:146–150
17. Sedlack RE (2007) Validation of computer simulation training for esophagogastroduodenoscopy: Pilot study. J Gastroenterol Hepatol 22:1214–1219
18. Spaelter U, Moix T, Ilic D, Bleuler H, Bajka M (2004) A 4-dof haptic device for hysteroscopy simulation. Proc Int Conf Intell Robots Syst 4:3257–3263
19. Sierra R, Bajka M, Székely G (2006) Tumor growth models to generate pathologies for surgical training simulators. Med Image Anal 10:305–316
20. Sierra R, Zsemlye G, Székely G, Bajka M (2006) Generation of variable anatomical models for surgical training simulators. Med Image Anal 10:275–285
21. Gallagher AG, Ritter EM, Champion H, G.Higgins, Fried M, Moses G (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 241:364–372
22. Tuchschmid S, Bajka M, Bachofen D, Székely G, Harders M (2007) Objective surgical performance assessment for virtual hysteroscopy. Stud Health Technol Inform 125:473–478
23. Pearlman MD (2006) Patient safety in obstetrics and gynecology: an agenda for the future. Obstet Gynecol 108(5):1266–1271
24. Fenner DE (2005) Training of a gynecologic surgeon. Obstet Gynecol 105:193–196