



Construct and face validity and task workload for laparoscopic camera navigation: virtual reality versus videotrainer systems at the SAGES Learning Center

Dimitrios Stefanidis,¹ Randy Haluck,² Tai Pham,² J. Bruce Dunne,³ Timothy Reinke,² Sarah Markley,³ James R. Korndorffer Jr.,³ Paul Arellano,⁴ Daniel B. Jones,⁵ Daniel J. Scott⁶

¹ Department of General Surgery - Division of Gastrointestinal and Minimally Invasive Surgery, Carolinas Medical Center, Charlotte, NC, USA

² Department of Surgery, Pennsylvania State University, Philadelphia, PA, USA

³ Tulane Center for Minimally Invasive Surgery, Tulane University Health Sciences Center, New Orleans, LA, USA

⁴ Texas Endosurgery Institute, San Antonio, TX, USA

⁵ Section of Minimally Invasive Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

⁶ Southwestern Center for Minimally Invasive Surgery, University of Texas Southwestern Medical Center, Dallas, TX, USA

Received: 18 July 2006/Accepted: 16 October 2006/Online publication: 6 December 2006

Abstract

Background: Laparoscopic camera navigation (LCN) training on simulators has demonstrated transferability to actual operations, but no comparative data exist. The objective of this study was to compare the construct and face validity, as well as workload, of two previously validated virtual reality (VR) and videotrainer (VT) systems.

Methods: Attendees ($n = 90$) of the SAGES 2005 Learning Center performed two repetitions on both VR (EndoTower™) and VT (Tulane Trainer) LCN systems using 30° laparoscopes and completed a questionnaire regarding demographics, simulator characteristics, and task workload. Construct validity was determined by comparing the performance scores of subjects with various levels of experience according to five parameters and face validity according to eight. The validated NASA-TLX questionnaire that rates the mental, physical, and temporal demand of a task as well as the performance, effort, and frustration of the subject was used for workload measurement.

Results: Construct validity was demonstrated for both simulators according to the number of basic laparoscopic cases ($p = 0.005$), number of advanced cases ($p < 0.001$), and frequency of angled scope use ($p < 0.001$), and only for VT according to training level ($p < 0.001$) and fellowship training ($p = 0.008$). Face validity ratings on a 1-20 scale averaged 15.4 ± 3 for VR vs. 16 ± 2.6 for VT ($p = 0.04$). Ninety-six percent

of participants rated both simulators as valid educational tools. The NASA-TLX overall workload score was 69.5 ± 24 for VR vs. 68.8 ± 20.5 for VT ($p = 0.31$).

Conclusions: This is the largest study to date that compares two validated LCN simulators. While subtle differences exist, both VR and VT simulators demonstrated excellent construct validity, good face validity, and acceptable workload parameters. These systems thus represent useful training devices and should be widely used to improve surgical performance.

Key words: Simulation — Skills assessment — Virtual reality — Laparoscopic camera navigation — Construct validity — Face validity — Workload

Laparoscopy has not only revolutionized surgery by improving patient outcomes but has also presented the surgeon with a new challenge by introducing a new skill set that needs to be mastered. Contrary to open surgery where surgeons have complete visual and tactile control of the operating field, during laparoscopy both those senses are altered. Tactile feedback from the operating field is severely limited due to the interposition of long instruments between the surgeon's hands and the tissues; the surgeons have to rely on visual cues from the operating field to compensate for the compromised haptics. Visual control, however, is also significantly impaired by the loss of depth perception (two-dimensional monitors) [10] and because the surgeon has to rely on an assistant who controls the camera and thus the field of view. Nevertheless, with appropriate training and experience,

surgeons adapt to these limitations of laparoscopy and learn how to use some laparoscopic features like the magnified view to their advantage.

Because the laparoscope is being controlled manually, its movement is impaired by some well-known obstacles of laparoscopic surgery such as the fulcrum effect, fixed access points, and decreased range of motion [12, 13]. Furthermore, many functions of the human eye such as target centering and sizing, focusing, tracking, steadiness, and maintenance of orientation have to be accomplished by the laparoscope and its operator [7]. Therefore, it is evident that proper visualization during a laparoscopic procedure depends on the navigation ability of the camera operator. Inadequate camera navigation ability may impair the surgeon's visualization, cause frustration, and compromise patient safety. The skill required, however, is far from being intuitive and camera operators may need to overcome a considerable learning curve [7]. Given that in most instances the camera operator is the least experienced member of the operating team, the importance of training in laparoscopic camera navigation (LCN) outside of the operating room cannot be overemphasized.

In an era where simulation is becoming an important component of surgical training, a number of simulators offer training modules for LCN. Nevertheless, only two simulators (EndoTower™, Verifi Technologies Inc., Elizabethtown, PA, and Tulane Trainer, Tulane Center for Minimally Invasive Surgery, New Orleans, LA) focus solely on LCN and have been validated as teaching tools for LCN [3, 6]. While these two simulators use fundamentally different operating platforms (virtual reality for EndoTower and video-trainer for Tulane Trainer), both have been shown to improve the operating room performance of trainees [2, 7]. Nevertheless, no comparative studies exist between these two simulators, especially using a large cohort of surgeons with variable laparoscopic expertise, and further external validation is needed. Furthermore, although workload can have a significant effect on learning, we know very little about the workload demands imposed on trainees by simulators [15].

The purpose of this study was to compare the EndoTower and Tulane Trainer LCN simulators for construct and face validity as well as operator workload. Moreover, we also aimed to externally validate these simulators in a large cohort of surgeons and obtain data from expert participants who might be used as a reference for assessment purposes.

Materials and methods

Attendees of the Laparoscopic Camera Navigation Station at the Learning Center (Fig. 1) during the 2005 Society of American Gastrointestinal Endoscopic Surgeons (SAGES) Annual Meeting in Hollywood, Florida, were eligible for participation. Study participation was voluntary; attendees who did not wish to be enrolled in the study and those who had prior experience with either simulator were excluded.



Fig. 1. Laparoscopic Camera Navigation Station at the SAGES 2005 Learning Center with participant on the EndoTower™ station.

Study design

Participants completed a short survey detailing demographic information, previous laparoscopic experience, and angled laparoscope use in their practice; to blind the instructors this information was hidden by folding the questionnaire. Participants then performed two repetitions on both a Virtual Reality Camera Navigation simulator (EndoTower) and a Laparoscopic Camera Navigation Videotrainer (Tulane Trainer) using the 30° angled laparoscope; their performances were recorded. Standardized instruction was given to all participants before initiation of the task by six experienced instructors and feedback was provided on an as-needed basis. The order of which simulator was used first was random and based on availability. Availability was maximized by applying a 5-minute cutoff time per repetition. Upon completion of the tasks, participants evaluated both simulators on a 20-point visual analog scale (VAS; from 1 = poor to 20 = excellent) with respect to their (1) value as training tools, (2) value as testing tools, (3) overall relevance to actual laparoscopic surgery, (4) ability to teach/test target centering and sizing, (5) ability to teach/test camera angle adjustment, (6) ability to teach/test horizon acquisition, (7) ability to teach/test camera steadiness, and (8) ability to provide performance feedback. In addition, participants were asked whether each simulator was a valid educational tool. Furthermore, the workload imposed on participants by each simulator was assessed on the same VAS scale using the NASA-TLX questionnaire [5]. This validated subjective workload assessment tool allows participants to evaluate the mental, physical, and temporal demands of the task and to report their effort, frustration, and perceived performance.

Previously validated proficiency scores for each simulator were made known to all participants before performing the tasks. These proficiency scores had been determined by averaging the scores of known experts during their performance of multiple repetitions of the same tasks. The proficiency score for the EndoTower task was 49 and for the LCN task 80 (low scores indicate superior performance) [7, 8]. Participants who had performed over 100 advanced laparoscopy cases and with an angled scope use of at least 50% frequency in their practice were considered experts in laparoscopic camera navigation. To eval-

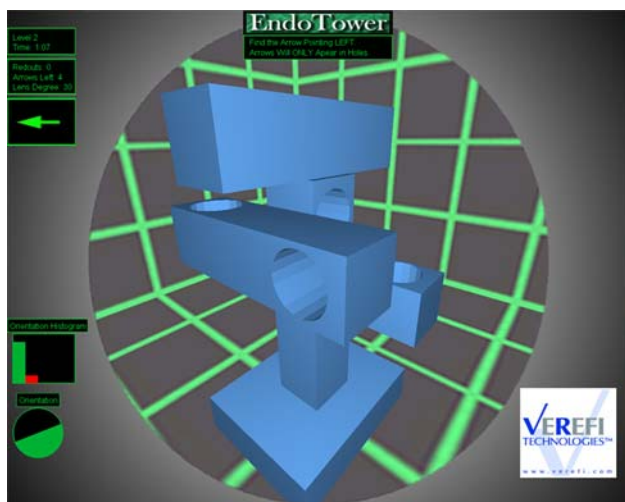


Fig. 2. The three-dimensional EndoTower™ simulator.

uate the presence of any simulator interface issues that might reduce the degree of fidelity of the simulated task to actual laparoscopic surgery by introducing a simulator-specific learning curve, expert participants' performance was compared to the proficiency scores. Furthermore, participant performance scores during the first and second repetitions on each simulator were compared to measure the simulator-associated learning curves.

Laparoscopic camera navigation simulators

EndoTower (Fig. 2)

The EndoTower consists of the Virtual Laparoscopic Interface (Immersion Corporation, San Jose, CA) used in conjunction with specially designed software (Verifi Technologies Inc., Elizabethtown, PA) and a slightly different handpiece that simulates the mechanical workings of the laparoscope [4]. The subject must navigate a virtual 30° laparoscope by manipulating the simulated camera handpiece that facilitates camera angle adjustment and rotational and overall camera positioning. The EndoTower system's virtual "tower" is a computer-generated, complex, three-dimensional abstract object that serves as an obstacle course that must be successfully traversed to find hidden objects; six arrows are randomly placed on and within seven target locations. Subjects have to acquire appropriate laparoscopic views of each arrow and hold the image steady for 2 s until the arrow disappears before proceeding to the next target. Any collision with the EndoTower causes the view to become blurry and red ("red out"), simulating touching the laparoscope to an organ and smudging the lens with blood. The learner must withdraw the virtual laparoscope and await a cleaning mode, thus penalizing the user's score in terms of time, efficiency, and errors. System metrics include (1) task completion time, (2) total path length (PL), (3) total rotational path length (RPL), and (4) percent of time off level (%TOL). An on-screen leveling gauge provides real-time feedback to the user. An overall score based on these metrics was calculated for each participant using the formula: $\text{Time} + [(PL/303) * 10] + [(RPL/22) * 10] + \%TOL$, where 303 and 22 represent expert performance scores for those individual parameters, respectively [7, 8]. The system offers three levels of task difficulty; level 1 was chosen for this study because this configuration generates the targets in the most standardized fashion.

Tulane Trainer (Fig. 3)

As previously described [7], this simulator is a patented yet not commercially available physical model used in conjunction with a videotrainer. The model consists of six 4-cm red circular targets housed in square open-faced white boxes with the edge of the box's face cut to

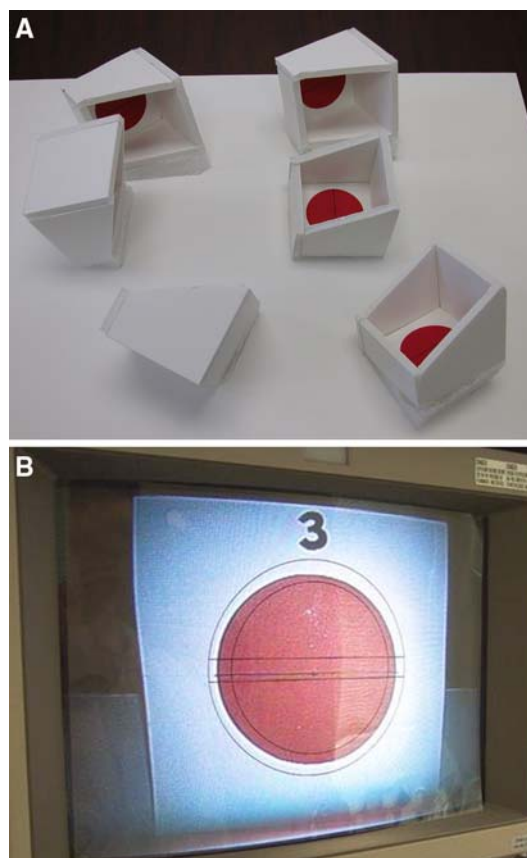


Fig. 3. Tulane Trainer: Left image depicts the six targets of the Tulane Trainer board and the right the monitor-mounted target acquisition template.

match the face angle of the laparoscope. The boxes are mounted on a board that is placed within a videotrainer such that all targets must be sequentially acquired by the learner with the 30° angled laparoscope in clockwise order. Each target is marked with a black line to indicate its true horizontal axis. A monitor-mounted target acquisition template (transparency sheet held by static electricity) containing two concentric circles and two parallel lines was used to verify appropriate sizing, centering, and horizontal orientation. Participants must maintain a correct image of each target within the monitor-mounted template for 5-s, as verified via direct observation by the instructor using a countdown timer with an audible alarm (Robic, Oxford, CT). If an error occurs by a drift outside the lines of the template, the 5-s countdown is restarted after the target is correctly reacquired. Total completion time is determined using a stopwatch (Fisher Scientific, Hampton, NH). Participant scores are based on total completion time which incorporates errors.

Statistical analysis

Statistical analysis was performed using the Sigma Stat 3.0 statistical software (SPSS Inc., Chicago, IL). Construct validity was assessed for each simulator by comparing scores of subjects stratified according to level of training, advanced laparoscopic fellowship training, experience with basic and advanced laparoscopic cases, and experience with an angled (30° or 45°) laparoscope using analysis of variance (ANOVA). Face validity and workload ratings of the simulators as well as performance scores during the first and second repetition were compared using paired *t* tests. Outliers were defined as those participants whose performance scores were outside two standard deviations (SD) of the group average. $p < 0.05$ was considered significant. Results are reported as mean \pm SD unless otherwise stated.

Table 1. Participant experience and construct validity of simulators

Parameter	<i>n</i>	Tulane Trainer		EndoTower™		
		Score ^a (mean ± SD)	<i>p</i> value	Score ^a (mean ± SD)	<i>p</i> value	
Training level	Novices	5	214 ± 58	<0.001	314 ± 166	0.17
	Residents	20	141 ± 91		222 ± 127	
	Fellows	15	108 ± 45		223 ± 150	
	Attendings	50	112 ± 50		203 ± 120	
Prior laparoscopic fellowship	No	56	130 ± 72	0.003	222 ± 142	0.49
	Yes	34	114 ± 53		215 ± 140	
Number laparoscopic cholecystectomies performed	None	4	273 ± 134	<0.001	306 ± 160	0.005
	1–0	16	154 ± 62		279 ± 154	
	51–100	16	119 ± 47		226 ± 144	
	101–250	19	108 ± 51		221 ± 139	
	> 251	35	103 ± 41		180 ± 97	
Number advanced laparoscopy cases performed	None	13	206 ± 100	<0.001	320 ± 175	<0.001
	1–5	22	130 ± 55		263 ± 144	
	26–50	13	121 ± 45		222 ± 141	
	51–100	8	115 ± 59		181 ± 77	
	> 101	34	91 ± 25		163 ± 73	
Frequency of angled scope use	Never	8	225 ± 115	<0.001	339 ± 166	<0.001
	< 50% of time	21	144 ± 66		282 ± 162	
	> 50% of time	61	103 ± 37		182 ± 95	

^a Low scores indicate superior performance

Results

One hundred thirteen attendees of the 2005 SAGES meeting participated in the activities of the Laparoscopic Camera Navigation Station of the Learning Center. Ninety (80%) participants who performed two repetitions on each simulator and fully completed the questionnaire were included in the study; subjects who had participated on only one simulator or for whom complete data were not available were excluded from further analysis. Mean participant age was 38 years (range = 25–66); 68 (76%) participants were men and 83 (92%) were right-hand dominant. Fifty (55%) subjects were practicing surgeons, 15 (17%) were advanced laparoscopy fellows, 20 (22%) were general surgery residents, and 5 (6%) were novices. The 50 surgeons were in practice for a mean of 9.7 ± 7.7 years and were practicing laparoscopy for an average of 7.5 ± 5.3 years; 16 (32%) of them had previously completed a formal advanced laparoscopy training fellowship. Prior laparoscopic experience and frequency of angled scope use in their practice is shown in Table 1.

Construct validity

The Tulane Trainer clearly distinguished participants according to prior fellowship training in laparoscopy, training level, number of laparoscopic cholecystectomies and advanced laparoscopic cases performed, and frequency of angled scope use in their practice (Table 1). On the other hand, the EndoTower did not distinguish participants according to fellowship training in laparoscopy or training level but clearly distinguished them according to the number of laparoscopic cholecystec-

tomies and advanced laparoscopic cases performed and frequency of angled scope use in their practice (Table 1).

Face validity

Both simulators were rated as valid educational tools by 96% of participants ($p = \text{n.s.}$). According to participant assessment on the 20-point VAS scales, no statistically significant differences were found between the two simulators in their validity as training and testing tools, in their relevance to actual laparoscopic surgery, in their ability to teach or test camera steadiness, and to provide performance feedback (Fig. 4). On the other hand, the Tulane Trainer achieved better ratings for its ability to teach or test target centering and sizing, camera angle adjustment, and horizon acquisition. Moreover, participants rated the feedback they received from instructors higher for the Tulane Trainer (16.8 ± 2.7 vs. 16.3 ± 3.1 for EndoTower; $p = 0.04$). Ratings in all categories and for both simulators were not influenced by prior experience with laparoscopy, because no statistically significant differences were found in the ratings of more experienced participants compared to the less experienced ones.

Workload assessment

With the exception of the frustration level, which was slightly higher for the EndoTower (9 ± 6 vs. 8 ± 5 for TT; $p = 0.04$), there were no statistically significant differences between the two simulators according to the mental, physical, and temporal demands of the task as well as the self-reported effort and performance of the participants as examined with the NASA-TLX workload assessment questionnaire (Fig. 5).

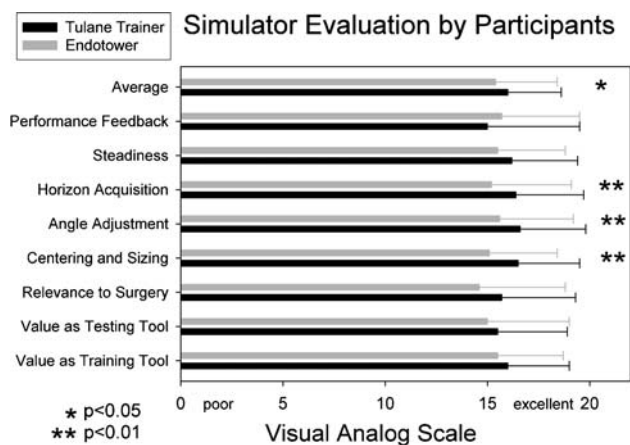


Fig. 4. Simulator Face Validity: Average ratings on eight face validity parameters along with standard deviations are depicted. *p* values reflect comparisons between simulators using the paired *t* test.

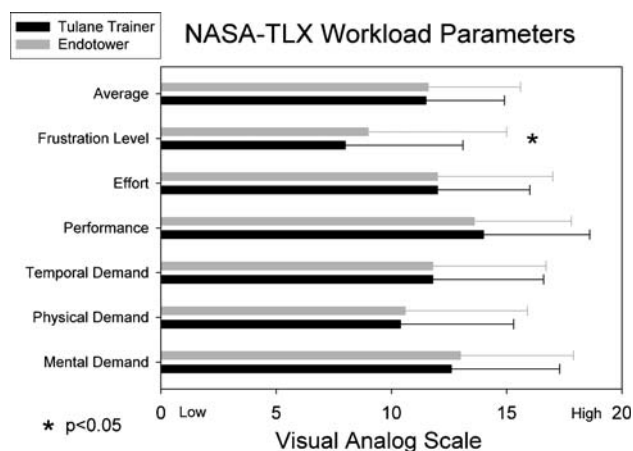


Fig. 5. Participant Workload: Average ratings on the six NASA-TLX workload parameters along with standard deviations are depicted. *p* values reflect comparisons between simulators using the paired *t* test.

Expert performance and simulator-associated learning curves

Thirty participants who had performed over 100 advanced laparoscopic procedures and were using an angled laparoscope in over one half of their cases met the criteria for inclusion in the expert group. Only five (17%) surgeons from this group achieved the proficiency score on the Tulane Trainer during their first repetition, whereas 19 (63%) achieved it during their second repetition. No expert participant achieved the proficiency score on the EndoTower. After outliers (two for Tulane Trainer and three for EndoTower) were excluded, the expert group performance was 95 ± 18 during the first repetition vs. 76 ± 9 during the second repetition for the Tulane Trainer ($p < 0.001$) and 161 ± 46 vs. 117 ± 29 for the EndoTower ($p < 0.001$), respectively.

The performance of all study participants improved on average by 23% on the Tulane Trainer (140 ± 70 vs. 107 ± 57 ; 23%; $p < 0.001$) and by 28% on the EndoTower (256 ± 145 vs. 184 ± 111 ; $p < 0.001$) between the first and second repetitions. Furthermore, 67% of

participants on the EndoTower and 91% on the Tulane Trainer demonstrated improved performance during the second attempt indicating a learning curve for both simulators.

Discussion

This is the largest study to date that compares two laparoscopic camera navigation simulators; it was conducted at the SAGES Learning Center, which allowed for the participation of a large cohort of surgeons with variable laparoscopic experience. Our findings demonstrate that both a virtual reality and a videotrainer LCN simulator possess excellent construct validity and good face validity and subject the learner to acceptable workload. Indeed, the Tulane Trainer demonstrated construct validity according to all five parameters examined and regardless of how experience with laparoscopy and laparoscopic camera navigation was categorized (training level, prior advanced laparoscopy fellowship, number of laparoscopic cholecystectomies performed, number of advanced laparoscopic cases performed, and frequency of angled scope use in practice). On the other hand, the EndoTower system distinguished participants according to three of the five parameters, not distinguishing subjects according to training level and prior fellowship training. Nonetheless, even though statistical significance was not demonstrated according to the latter two parameters, more experienced participants outperformed those with less laparoscopic experience (Table 1). Hence, our sample size may not have been large enough to detect these smaller differences. In addition, the parameters that did reveal construct validity for the EndoTower are probably the most important with respect to laparoscopic camera navigation ability; the number of advanced laparoscopic cases performed (during which an angled scope is usually being used) and the frequency of angled scope use in daily practice likely reflect much better true camera navigation ability than the other parameters examined. For both of these categories of laparoscopic experience, both simulators demonstrated construct validity at highly significant statistical levels ($p < 0.001$). Construct validity for the simulators used in this study has been previously shown in limited studies with small numbers of participants [2, 4, 6]. The construct validity of EndoTower was also assessed in a prior study from the 2004 SAGES Learning Center [8]. In a surgeon cohort very similar to our study, the EndoTower was found to distinguish participants according to prior fellowship training and advanced laparoscopic experience but not according to basic laparoscopic experience or training level. Methodology differences between that study and our study may account for subtle discrepancies in the findings.

Face validity represents the extent of realism of the simulation to the actual task that is simulated [12]. Both simulators in this study were given very good face validity ratings by participants; the Tulane Trainer averaged 16 ± 2.6 and the EndoTower 15.4 ± 3 on a

20-point VAS scale ($p = 0.04$). Interface issues may account for the face validity differences we encountered between the two systems, with the more realistic Tulane Trainer achieving higher ratings. As has been suggested before, the perception of the more “game-like” environment of a virtual reality system may be less realistic compared with a box-trainer model, where subjects use the same instruments and equipment as in the operating room [8, 14]. On the other hand, the differences in centering, sizing, and horizon acquisition ability may have been also a consequence of dissimilar simulator designs; while the Tulane Trainer was designed to allow for up to only a 24% size change and 9° horizontal drift during correct target acquisition [7], targets on the EndoTower can be successfully acquired while off center and only partially in view because this is thought to benefit learning [4]. Nevertheless, the absolute difference in face validity ratings of the two simulators in this study (0.6 on a 20-point scale or 3%) is negligible. From this perspective, the EndoTower achieved face validity ratings almost identical to a realistic trainer indicating the quality of the EndoTower virtual reality platform. Very similar face validity ratings for the EndoTower were also documented in the study by Maithel et al. [8]; they documented an average rating of 7.9 on a 1-10 scale of six slightly different face validity parameters. If the average face validity ratings of the Maithel et al. study and our study are normalized to the maximum value of the respective rating scales (10 and 20, respectively) to allow direct comparison, the values obtained reveal very good and essentially identical face validity ratings (79% vs. 77%, respectively; $p = \text{n.s.}$) in both studies for the EndoTower. This finding supports the reliability of this simulator.

Surprisingly, despite the ability of the virtual reality simulator to provide additional performance metrics and “red-outs,” it did not achieve significantly higher ratings in its ability to provide performance feedback to the learner. The presence of instructors for both simulators, however, may have influenced these ratings; while instructors had been advised to provide minimal feedback during the performance of the tasks, feedback was given when deemed necessary or requested by the participants. The Tulane Trainer requires, however, that the proctor alert the participant to “drifts” outside of the target acquisition template, whereas the EndoTower feedback is generated primarily by the computer. Although we did not measure instructor feedback directly, the participants were asked to evaluate the feedback they received by the instructors. Participant ratings indicated that instructor feedback was slightly more on the Tulane Trainer (16.8 ± 2.7 vs. 16.3 ± 3.1 for EndoTower; $p = 0.04$), which could have influenced their ratings for simulator feedback. Nevertheless, this difference (3%) is again too small to be of any practical importance.

Workload is the quantity of a person’s cognitive capacity necessary to perform a certain task [9]. Its assessment is important for a more global evaluation of performance because traditionally used performance metrics do not account for the effort of the learner. To better assess participant performance on the simulators,

we therefore included the validated NASA-TLX workload assessment questionnaire in the evaluations of our simulators. Overall scores revealed moderate workload for both simulators without any significant differences (68.8 ± 20.5 for Tulane Trainer vs. 69.5 ± 24 for EndoTower; $p = \text{n.s.}$), except for a higher frustration level with the EndoTower. The 5% higher frustration rating of the EndoTower may be a consequence of the less realistic virtual reality interface, but it represents too small a difference to be of any practical significance. Overall, both simulators were associated with moderate participant effort, which may be ideal for training; the tasks were neither too easy nor too hard, conditions that may lead to early loss of interest in a simulator. Such workload characteristics may promote higher adaptation rates of these simulators by learners.

Expertise is domain specific and is not well studied in medicine [1]. Expert performance is characterized by consistency and maximal adaptation to task constraints [1]. Along those lines, highly experienced and capable laparoscopists (“experts”) should perform a simulated laparoscopic task with proficiency and consistency. On the other hand, if expert performance on a simulator produces inconsistent results during multiple attempts, it may point toward a simulator-specific learning curve (interface) that is unrelated to learning the primary task. In the context of our study, experts in laparoscopic camera navigation should not demonstrate performance differences between the first and second repetitions on each simulator. Because expert ability is unlikely to change from one repetition to the next, any performance changes during multiple repetitions on a simulator are likely reflective of simulator interface issues. In other words, experts learn how to use a simulator they are unfamiliar with rather than improving their ability. Furthermore, the longer it takes experts unfamiliar with a simulator to achieve the proficiency levels (established by experts familiar with the simulator), the longer the simulator-specific learning curve. This is an important simulator characteristic that has implications for training; novices using these devices may not only have to learn the simulated task but to also adjust to the interface of the simulator. The interface also has important implications for assessment, because the methodology for use of simulators in assuring competency is not yet well developed, except for a single program validated for high-stakes assessment [11].

To investigate such interface issues for the Tulane Trainer and EndoTower simulators, we examined the performance of the most experienced participant subgroup in LCN. In this “LCN expert” group we included participants who had performed over 100 advanced laparoscopic cases and who had frequently used (in over 50% of their cases) an angled scope in their practice. While these criteria are arbitrary, they were chosen because they selected the most experienced participants in LCN. Expert participants demonstrated significant performance improvement on the second repetition compared with the first on both simulators which points toward a simulator-specific learning curve. While 63% of expert participants achieved the proficiency level during their second repetition on the Tulane Trainer, no par-

participant achieved it on the EndoTower. In the study by Maithel et al. [8], the same observation was made: no participant was able to achieve the EndoTower proficiency level with two repetitions. This finding may well reflect more interface issues of this virtual reality simulator compared with the videotrainer model. Besides interface differences, however, one explanation for this discrepancy between the virtual reality and videotrainer LCN simulators could also be a consequence of our methodology; while known experts followed the most efficient order of target acquisition on both simulators for the development of proficiency levels, study participants were instructed to follow the same target order only on the Tulane Trainer. In contrast, the order of target acquisition on the EndoTower was random, as suggested by the manufacturer, and may have impaired participant efficiency and led to worse scores. Nevertheless, the random order acquisition was chosen because it was felt to be an important characteristic of the EndoTower that enhances trainee learning. On the other hand, the two simulators may have similar simulator-associated learning curves because the improvement in expert participant performance between the first and second repetitions was similar for both (20% for Tulane Trainer vs. 27% for EndoTower; $p = \text{n.s.}$). We therefore think that the proficiency level we used for the EndoTower may have been too stringent and improbable to achieve when targets are acquired without a prespecified order; as a result, the performance scores achieved by our expert participants likely provide a more accurate reflection of proficiency under these circumstances. Thus, our study offers expert data for laparoscopic camera navigation that can be used as reference points (i.e., proficiency levels) if these simulators were to be used as assessment tools. Because our expert participants were using the simulators for the first time and may not have overcome the simulator-associated learning curves after just two repetitions, these levels may be less suitable as training goals. Nevertheless, during their second repetition on the Tulane Trainer, 63% of expert participants achieved our proficiency level (80) and their average performance did not differ from our proficiency level (76 vs. 80). This provides evidence that most participants had overcome the Tulane Trainer interface by their second repetition and provides further external validation for our proficiency level, which has been demonstrated to improve trainee operative performance in the context of a proficiency-based curriculum [7].

The findings of our study demonstrate the utility of both simulators. It provides further external validation in a large cohort of surgeons of two previously validated LCN simulators and has implications for their use for skill assessment and training purposes. Ultimately, the vast majority of participants (96%) of our study rated both simulators as valid educational tools.

In conclusion, while subtle differences exist, both the Tulane Trainer and the EndoTower demonstrated

excellent construct validity, good face validity, and acceptable workload parameters in a large surgeon cohort. These systems represent useful training devices for laparoscopic camera navigation and should be widely used to improve surgical performance.

Acknowledgments. The authors gratefully acknowledge SAGES, Verifi Inc., and Karl Storz Endoscopy for equipment support.

References

- Ericsson KA (2004) Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 79: S70–S81
- Ganai S, Seymour NE (2005) VR to OR for Camera Navigation. In: Westwood JD, et al. (eds), *Medicine Meets Virtual Reality 13*, vol 111, IOC press, Amsterdam, pp 45–48
- Haluck RS, Gallagher AG, Satava RM, Webster R, Bass TL, Miller CA (2002) Reliability and validity of Endotower, a virtual reality trainer for angled endoscope navigation. *Stud Health Technol Inform* 85: 179–184
- Haluck RS, Webster RW, Snyder AJ, Melkonian MG, Mohler BJ, Dise ML, Lefever A (2001) A virtual reality surgical trainer for navigation in laparoscopic surgery. *Stud Health Technol Inform* 81: 171–176
- Hart SG, Staveland LE (1988) Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds), *Human Mental Workload*. Elsevier, Amsterdam, pp 139–183
- Korndorffer JR Jr, Stefanidis D, Sierra R, Clayton JL, et al. (2005) Validity and reliability of a videotrainer laparoscopic camera navigation simulator. *Surg Endosc* 19(Suppl): S246
- Korndorffer JR Jr, Hayes DJ, Dunne JB, Sierra R, Touchard CL, Markert RJ, Scott DJ (2005) Development and transferability of a cost-effective laparoscopic camera navigation simulator. *Surg Endosc* 19: 161–167
- Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D (2006) Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? *Surg Endosc* 20: 104–112
- O'Donnell RD, Eggemeier FT (1986) Workload assessment methodology. In: Boff KR, Kaufman L, Thomas JP (eds), *Handbook of perception and performance*, vol. 2: Cognitive processes and performance Wiley, New York, pp 42.1–42.49
- Perkins N, Starkes JL, Lee TD, Hutchison C (2002) Learning to use minimal access surgical instruments and 2-dimensional remote visual feedback: how difficult is the task for novices? *Adv Health Sci Educ Theory Pract* 7: 117–131
- Peters JH, Fried GM, Swanstrom LL, Soper NJ, Sillin LF, Schirmer B, Foffman K, SAGES FLS Committee (2004) Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery* 135: 21–27
- Scott DJ, Jones DB (2005) Virtual reality training and teaching tools. In: Soper NJ, Swanstrom LL, Eubanks WS (eds), *Mastery of Endoscopic and Laparoscopic Surgery*, 2nd ed. Lippincott Williams & Wilkins, Philadelphia, pp 146–160
- Stefanidis D, Korndorffer JR, Scott DJ (2005) Robotic laparoscopic fundoplication. *Curr Treat Options Gastroenterol* 8: 71–83
- Stefanidis D, Korndorffer JR Jr, Sierra R, Touchard C, Dunne JB, Scott DJ (2005) Skill retention following proficiency-based laparoscopic simulator training. *Surgery* 138: 165–170
- Wickens CD, Hollands J (2000) *Engineering psychology and human performance*. Prentice Hall, Upper Saddle River, NJ