Surg Endosc (2008) 22:2621–2624
DOI 10.1007/s00464-008-9844-9

# Time-efficient laparoscopic skills assessment using an augmented-reality simulator

**J. Adam Oostema · Matthew P. Abdel · Jon C. Gould**

Received: 29 May 2007 / Accepted: 28 November 2007 / Published online: 18 March 2008


**Abstract**
*Background* Computer-based, virtual-reality laparoscopic surgical simulators have several advantages over traditional video trainers. One of these advantages is that performance can be evaluated using unique computer-derived metrics, which can be digitally archived for analysis at a time convenient to instructors. This study sought to determine whether the computer-derived metrics for a unique hybrid simulator correlated with laparoscopic surgical skill.
*Methods* For this study, 24 medical students (3rd year), 19 surgical residents (postgraduate years 1–5), and 3 attending surgeons were invited to perform four different tasks three times in a hybrid laparoscopic trainer (ProMIS). Instruction with minimal supervision occurred at a time convenient to each subject. The four tasks in order of complexity were laparoscopic orientation, object positioning, sharp dissection, and intracorporeal knot tying. The metrics automatically recorded were time, path length, and smoothness. The laparoscopic operative experience for each user was quantified using case logs.
*Results* A statistically significant correlation was observed between experience and performance for all three metrics for tasks 2 to 4 ($p < 0.01$). Smoothness was the only metric that correlated with the laparoscopic orientation task. Within tasks, time and smoothness correlated much more strongly with experience and to a similar degree. The strongest correlation was observed for the knot-tying task ($r^2 = 0.60$ for time and 0.59 smoothness).
*Conclusions* The computer-derived metrics measured by the hybrid trainer correlate with laparoscopic experience. These metrics are automatically calculated and stored. This may make skills assessment and training a more time-efficient endeavor for instructors and trainees alike. Further study is necessary to determine whether specific metrics are better indicators of actual skill.

**Keywords** Education · Training/courses

J. A. Oostema · M. P. Abdel · J. C. Gould (✉)
Department of Surgery, University of Wisconsin School of Medicine and Public Health, 600 Highland Avenue K4/762, Clinical Science Center, Madison, WI 53792, USA
e-mail: gould@surgery.wisc.edu

In the early days of laparoscopic surgery, surgeons and residents alike often developed and refined their skills in the operating room. Recent pressures such as the 80-h workweek, a focus on medical errors, and the increasing cost of operating room time have led surgeons and educators to develop tools for enhancing surgical skills in safe, reproducible, and measurable environments. Surgical skills curricula using video trainers meet these criteria. Video trainers are inexpensive, portable, and readily available. Several video trainer–based laparoscopic skills curricula have been validated and are in widespread use [1–3].

Quantifying skill with a video trainer typically is limited to basic metrics such as time required to complete a task, penalties for committing predefined errors, or both. Video trainer–based assessments are further limited by the time commitment of instructors, who often are busy practicing surgeons, to administer the curriculum and evaluate a trainee's performance.

Recently, a new type of device has become commercially available that addresses several of these shortcomings. The virtual-reality trainer is capable of more realistic task simulations than can be accomplished with video box trainers. Many virtual-reality trainers digitally archive task

repetitions for viewing at a time convenient to instructors. Virtual-reality simulators are capable of assessing performance using computer-derived metrics, which are not attainable with traditional video trainers. Many of these computer-derived metrics have not yet been validated. A simulated task is said to possess evidence of construct validity if users with more experience, and by inference, more skill, perform better or more efficiently than users with less experience.

We aimed to determine whether computer-derived performance metrics for the ProMIS hybrid virtual reality trainer (Haptica Inc., Dublin, Ireland) correlate with prior laparoscopic experience. We also sought to determine whether these data could be collected in a time-efficient and convenient manner for instructors and trainees.

## Methods

For this study, 24 medical students (3rd year), 19 general surgery residents (postgraduate year 1–5), and 3 attending surgeons chose to participate and were asked to perform three repetitions of four tasks in the ProMIS hybrid virtual-reality trainer. The ProMIS trainer is similar to a traditional video trainer in that it has an actual physical space in which trainees can perform a simulated task using actual instruments with real haptic feedback. A trainee can interact with a computer-generated simulated environment or a real video image of the physical space within the trainer box, depending on the task. For all tasks, the computer tracks the movements of instruments in space and derives some unique performance metrics.

Tasks were selected on the basis of perceived complexity, from very easy to very difficult. These tasks, in order of increasing complexity, were camera navigation, object positioning, sharp dissection, and intracorporeal knot tying. In performing the camera navigation task, trainees maneuvered a laparoscope about a virtual abdomen and zoomed in on target objects. The object-positioning task involved picking up small beads from a central canister and transferring them to one of four numbered canisters with either the left or right hand as directed. The sharp dissection task required trainees to cut a circle from a surgical glove mounted on a task template. In the intracorporeal knot-tying task, the user was directed to tie a squared surgeon's knot using intracorporeal knot-tying techniques.

Performance metrics generated by the computer included time required to complete each task, smoothness, and path length. Path length was defined as the distance in millimeters that the tip of each instrument travels during a task repetition. Smoothness was defined as the number of times an instrument changes velocity during the performance of a task.

**Table 1** Experience score including basic and advanced scores

| Basic laparoscopic cases (*n*) | Advanced laparoscopic cases (*n*) |
|---|---|
| 0–5 = 0 | 0–5 = 0 |
| 6–15 = 1 | 6–15 = 2 |
| 16–50 = 2 | 16–50 = 4 |
| 51–100 = 3 | 51–100 = 6 |
| >100 = 4 | >100 = 8 |

Prior laparoscopic experience was quantified using updated operative case logs for each resident at the time of testing. The medical students had no prior operative experience. All three attending surgeons had performed approximately 1,000 or more laparoscopic procedures at the time of testing. On the basis of prior experience, an experience score was attained using a scoring system prospectively developed for this study (Table 1).

Many of the study data were collected without an instructor present. Participants were directed to watch the demonstration videos that accompany the specific tasks on the trainer's computer. All task repetitions were automatically archived to the hard drive of the computer for viewing at a time convenient to the instructor. Only metrics for correctly performed and completed tasks were included in the study data. For incomplete or incorrect task performance, the participants were asked to repeat the tasks in question until the metrics for three valid repetitions were recorded for each task.

Correlations were determined using regression analysis and analysis of variance (ANOVA). All statistical calculations were performed using Mstat v4.01 software (University of Wisconsin Medical School, Madison, WI, USA).

## Results

Experience scores correlated highly with postgraduate level of training ($r^2 = 0.92$; $p < 0.001$) (Table 2). For the subjectively easiest task (camera navigation), the only metric of the three recorded that correlated with experience was smoothness. For the other three increasingly complex tasks, all three metrics correlated with experience to a statistically significant degree. The correlations were the strongest for the subjectively most difficult task (intracorporeal knot tying). Within tasks, smoothness and time correlated more strongly with experience than did path length (Table 3).

**Table 2** Experience score vs level of training

| Training level | Mean calculated experience score |
|---|---|
| MS3 | 0 |
| PGY1 | 1.6 |
| PGY2 | 3.8 |
| PGY3 | 5.8 |
| PGY4 | 7.2 |
| PGY5 | 9.0 |
| Attending surgeon (−) | 12.0 |

MS3, third-year medical student; PGY, postgraduate year

**Table 3** Correlations between experience score and computer-derived metrics for selected tasks

| Task | Metric | Correlation ($r^2$) | $p$ value[a] |
|---|---|---|---|
| Camera navigation | Time | 0.07 | 0.09 |
| | Path length | 0.0003 | 0.9 |
| | Smoothness | 0.11 | 0.03 |
| Object positioning | Time | 0.36 | <0.01 |
| | Path length | 0.17 | 0.01 |
| | Smoothness | 0.36 | <0.01 |
| Sharp dissection | Time | 0.39 | <0.01 |
| | Path length | 0.33 | <0.01 |
| | Smoothness | 0.45 | <0.01 |
| Intracorporeal knot | Time | 0.60 | <0.01 |
| | Path length | 0.39 | <0.01 |
| | Smoothness | 0.59 | <0.01 |

[a] $p < 0.05$ considered significant

## Discussion

Recent pressures such as the 80-h workweek and the American College of Graduate Medical Education (ACGME) directive regarding general competencies challenge the traditional model of surgical training. The operating room is not the safest or necessarily even the most effective place to develop basic skills.

Fitts and Posner's [4] three-stage theory of motor skills acquisition is widely accepted in both the motor and surgical skills literatures. In the initial cognitive stage of learning a motor skill, performance is erratic, and basic tasks are broken down into distinct steps. In the operating room, this often results in the inefficient performance of a task, which may result in the conclusion of the educational opportunity as the attending takes over the case.

It has become clear that a basic set of laparoscopic surgical skills can be developed outside the operating room using simulators [5]. Furthermore, it has been demonstrated that these skills developed in a dry lab setting are real skills transferable to the operating room [6]. This is rapidly becoming the new paradigm of surgical education.

Unfortunately, training surgeons and residents in a skills lab can be costly and time consuming, on the part of both instructors and learners. A recent survey of American surgery residency program directors showed that the faculty and instructor time commitment for operating a surgical skills lab was the biggest obstacle encountered [7].

In the current study, we demonstrated that computer-based laparoscopic simulators have certain advantages over traditional video trainers that can help address some of these issues. The specific simulator used in this study is very user friendly and intuitive. Written descriptions and video demonstrations accompany all tasks. During task performance, audio instructions are provided for each step. Metric feedback is immediate, and video replays of task performance can be viewed by the learner with computer enhancements when desired. These features may facilitate learning at a time convenient for the trainee without necessarily requiring the presence of an instructor to apply the metrics or provide feedback. Feedback is still possible through video review at a time convenient to the instructor. Computer-derived metrics are recorded and can be exported as a data file for further analysis. Most of our study data were collected without an instructor present.

A laparoscopic skills curriculum based on computerized simulators such as the one used in this study could be implemented in a single introductory group session. Supplemental written and computer-based instructions would help provide direction during independent learning and training sessions. Time spent during each training session could be calculated and recorded for each trainee. Using a remote network, it is possible for an instructor to log-on and download performance metrics and video files for review. This would allow an instructor to monitor a trainee's progress, objectively measure surgical skills, and provide feedback even from a distance, such as when a resident travels for an outside rotation.

The time efficiency of training on a computer-based simulator is an important factor that ultimately may favor these kinds of trainers over traditional video trainers. Other authors have reached similar conclusions regarding virtual-reality simulators [8].

This study also provided some preliminary evidence that the computer-derived metrics of the ProMIS system possess construct validity and may be useful for measuring laparoscopic surgical skill. Several other recently published studies have suggested the same. Van Sickle et al. [9] demonstrated that the ProMIS metrics are capable of discriminating between expert and novice laparoscopists in the intracorporeal knot-tying task. Broe et al. [10] demonstrated that three selected simulated ProMIS tasks

are construct valid for differentiating experience levels among surgeons in training.

Further study is necessary to determine whether specific metrics are more accurate measures of actual skill. Our data suggest that time and smoothness are more sensitive measures of skill than path length. A validated scoring system and benchmarks could be established to guide training to defined proficiency levels on this device. This would be of great value to busy educators with an interest in efficient and effective assessment of trainees' laparoscopic surgical skills without major time commitments or sacrifices.

## References


1. Rosser JC, Rosser LE, Savalgi RS (1997) Skill acquisition and assessment for laparoscopic surgery. Arch Surg 132:200–204
2. Scott DJ, Young WN, Tesfay ST, Frawley WH, Rege RV, Jones DB (2001) Laparoscopic skills training. Am J Surg 182:137–142
3. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL (1998) Development of a model for training and evaluation of laparoscopic skills. Am J Surg 175:482–487
4. Fitts PM, Posner MI (1967) Human performance. Brooks/Cole, Belmont, CA
5. Fraser SA, Feldman LS, Stanbridge D, Fried GM (2005) Characterizing the learning curve for a basic laparoscopic drill. Surg Endosc 19:1572–1578
6. Fried GM, Feldman LS, Vassiliou MC, Fraser SA, Stanbridge D, Ghitulescu G, Andrew CG (2004) Proving the value of simulation in laparoscopic surgery. Ann Surg 240:518–525
7. Gould JC (2006) Building a laparoscopic skills lab: resources and support. JSLS 10:293–296
8. McDougall EM, Corica FA, Boker JR, Sala LG, Stoliar GS, Borin JF, Chu FT, Clayman RV (2006) Construct validity testing of a laparoscopic surgical simulator. J Am Coll Surg 202:779–787
9. Van Sickle KR, McClusky DA, Gallagher AG, Smith CD (2005) Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. Surg Endosc 19:1227–1231
10. Broe D, Ridgway PF, Johnson S, Tierney S, Conlon KC (2006) Construct validation of a novel hybrid surgical simulator. Surg Endosc 20:900–904