Skill retention following proficiency-based laparoscopic simulator training

Dimitrios Stefanidis, MD, PhD, James R. Korndorffer Jr, MD, Rafael Sierra, MD, Cheri Touchard, BS, J. Bruce Dunne, PhD, and Daniel J. Scott, MD, New Orleans, La

Background. Proficiency-based curricula using both virtual reality (VR) and videotrainer (VT) simulators have proven to be efficient and maximally effective, but little is known about the retention of acquired skills. The purpose of this study was to assess skill retention after completion of a validated laparoscopic skills curriculum.

Methods. Surgery residents (n = 14) with no previous VR or VT experience were enrolled in an Institutional Review Board–approved protocol and sequentially practiced 12 Minimally Invasive Surgical Trainer-VR and 5 VT tasks until proficiency levels were achieved. One VR (manipulate diathermy) and 1 VT (bean drop) tasks were selected for assessment at baseline, after training completion (posttest), and at retention.

Results. All residents completed the curriculum. Posttest assessment occurred at 13.2 ± 11.8 days and retention assessment at 7.0 ± 4.0 months. After an early performance decrement at posttest (17%-45%), the acquired skill was maintained up to the end of the follow-up period. For VR, scores were 81.5 ± 23.5 at baseline, 33.3 ± 1.8 at proficiency, 48.4 ± 9.2 at posttest, and 48.4 ± 11.8 at retention. For VT, scores were 49.4 ± 12.5 at baseline, 22.0 ± 1.4 at proficiency, 25.6 ± 3.6 at posttest, and 26.4 ± 4.2 at retention. Skill retention was better for VT, compared with VR (P < .02). The extent of skill deterioration did not correlate with training duration or resident level.

Conclusions. Although residents do not retain all acquired skills (more so for VR than for VT) according to simulator assessment, proficiency-based training on simulators results in durable skills. Additional studies are warranted to further optimize curriculum design, investigate simulator differences, and establish training methods that improve skill retention. (Surgery 2005;138:165-70.)

The teaching of operative skills in the clinical setting is constrained by the complexity of procedures, medicolegal and ethical concerns, fiscal and time limitations (especially in the 80-hour workweek), and has created the need for formal training outside the operating room.1-3 Moreover, the increased incidence of complications that was observed as a result of the undisciplined introduction of laparoscopic techniques in the early 1990s4 has raised public awareness and resulted in an outcry for safety.5 Loss of depth perception6 and haptic feedback, the fulcrum effect, and the use of instruments with limited range of motion7 make laparoscopic tasks difficult and introduce new skill sets that must be mastered. Acquisition of laparoscopic skills using both virtual reality (VR) and videotrainer (VT) simulators can help overcome the learning curve of new, complex, and difficult tasks, and lead to improved operative performance.8-12 Moreover, proficiency-based curricula have proven to be maximally effective and efficient.12-14 Besides skill acquisition, skill retention is vitally important15,16 but has not been well studied for surgical motor tasks. Contrary to the nonsurgical literature, in which many publications have investigated skill retention,15,17-24 only a handful of surgical papers have addressed this issue.25-27 In addition, there are no studies that compare the durability of skill between VR and VT simulators, which have previously demonstrated distinct learning characteristics.28 The purpose of this study was to investigate the durability of skill that surgery
residents acquired during a proficiency-based basic laparoscopic skills curriculum using both VR and VT simulators.

METHODS

Surgery residents (n = 14) of varying levels (R1-4) with no or minimal prior VR or VT experience were enrolled in an Institutional Review Board-approved training curriculum on 12 MIST-VR and five VT tasks. The Minimally Invasive Surgical Trainer (MIST)-VR (Mentice, Göteborg, Sweden) tasks consisted of 6 core skills 1 (CS1) and 6 core skills 2 (CS2) tasks on the easy default setting and the 5 Southwestern VT tasks of the bean drop, running string, checkerboard, block move, and suture foam drills using a 6-station VT (Karl-Storz Endoscopy, Culver City, Calif), as described in detail elsewhere.29 Scores for the MIST-VR were generated automatically by its software on the basis of completion time, errors, economy of movement, and economy of diathermy (for tasks using diathermy). VT scores were based on completion time and were recorded with a stopwatch (Fischer Scientific International Inc, Hampton, NH).

All residents completed a questionnaire regarding demographics, handedness, curriculum expectations, and prior experience with laparoscopic surgery, VR or VT simulators, and video games; responses were recorded on Likert scales (Table I). After a standardized demonstration of each task and administration of a baseline test at the beginning of the curriculum (3 repetitions of each task), all residents sequentially practiced the 17 tasks until previously established proficiency levels13,14 were achieved on 2 consecutive repetitions. The proficiency scores were posted at each station so that trainees could follow easily their own progress. Practice was scheduled during 1-hour weekly sessions, and a research assistant was available for assistance although no active instruction was given.

For the evaluation of skill retention, we elected to test trainees on 2 tasks to simplify the testing sessions. We chose the manipulate diathermy (MD) VR task (proficiency score 36.6) and the bean drop (BD) VT task (proficiency score 24) as these represent the most valid tasks of the 2 simulators.12,30

All residents performed 3 repetitions of both tasks after achieving proficiency (Posttest) and at the end of the academic year (Retention test). During the follow-up period, no resident had additional simulator exposure, but all had routine on-the-job training. Composite scores consisting of the mean of the 3 repetitions were calculated and used to compare performance differences. To compare relative performance on each simulator, we normalized composite scores according to the respective proficiency levels.

To examine expert performance over time, we tested 3 experts with extensive laparoscopic and simulator experience on both tasks at a time remote to any practice on the simulators and without allowing any warm up period on the tasks, thus mimicking the testing conditions for the subjects of this study. One of the experts also had been involved in the creation of the proficiency levels for both tasks.

Statistical analysis was performed with the use of a t test, paired t test, and Pearson’s correlation;

| Table I. Likert scales and distribution of responses in baseline questionnaire |
|-----------------------------|-----------------------------|
| Laparoscopy self-rating  | Course expectations  |
| 1. Very poor 7.7% | 1. Useful 69.2% |
| 2. Poor 46.1% | 2. Somewhat useful 30.8% |
| 3. Moderate 38.5% | 3. Not useful 0% |
| 4. Good 7.7% | |
| 5. Excellent 0% | |
| Past video game exposure | Current video game exposure |
| 1. None 7.7% | 1. None 69.3% |
| 2. Very little 30.7% | 2. Very little 23.0% |
| 3. Moderate 46.2% | 3. Moderate 7.7% |
| 4. Extensive 15.4% | 4. Extensive 0% |
| Prior VR experience | Prior VT experience |
| 1. None 84.6% | 1. None 84.6% |
| 2. <30 minutes 15.4% | 2. <30 minutes 15.4% |
| 3. <1 hour 0% | 3. <1 hour 0% |
| 4. <2 hours 0% | 4. <2 hours 0% |
| 5. <3 hours 0% | 5. <3 hours 0% |
RESULTS

All 14 residents successfully completed the curriculum; six R1, four R2, two R3, and two R4 residents participated. Mean age was 29.9 ± 3.1 years; 12 residents were male and 2, female; 12 residents were right handed, one, left handed, and one used both hands equally. Questionnaire data are shown in Table I.

Training on the MD and the BD tasks required 22.3 ± 15.3 and 19.9 ± 9.1 (P = NS) repetitions, respectively. The interval between achieving proficiency and posttest was 13.2 ± 11.8 days, and between proficiency and retention, 7.0 ± 4.0 months.

As seen in Figure 1, after an early performance decrement at posttest, there was no skill loss for both tasks at retention. In particular, for MD, scores were 81.5 ± 23.5 at baseline, 33.3 ± 1.8 at proficiency (59% improvement, compared with baseline, P < .001), 48.4 ± 9.2 at posttest (55% skill retention, compared with proficiency, P < .001), and 48.4 ± 11.8 at retention (0% deterioration, compared with posttest, P = NS). The early performance decay was greater for MD, compared with BD (45% vs 17%; P < .001). As seen in Figure 2, most of the skill loss occurred within the first 2 weeks after training and stabilized thereafter for the duration of follow-up. In contrast, as expected, the mean performance scores of the 3 experts on both tasks at a time remote to any simulator practice (37.0 for MD and 21 for BD) was very close to our proficiency levels (36.6 for MD and 24 for BD). Since only 1 of the 3 experts had been involved originally in the creation of our proficiency levels, we specifically examined his performance and found no changes over time (35.9 ± 2.1 at retest vs 36.6 ± 5.2 at level development for MD and 24.6 ± 3.2 vs 22.6 ± 5.4 for BD, respectively; P = NS for both).

There was no correlation of skill loss with resident level, duration of training (number of repetitions), or any of the other demographic and questionnaire parameters. Resident level and past laparoscopic experience correlated with baseline performance (indicating construct validity) but only for the BD (r = 0.68-0.71, P < .01). As seen in Figure 3, there was no difference in the performance between junior and senior residents for the MD task at any time interval: The BD baseline performance of seniors was superior to that of juniors (indicating further construct validity).
DISCUSSION

As seen in previous studies, our trainees achieved significant improvement in performance after completing a carefully structured proficiency-based curriculum on previously validated VR and VT simulators; skill acquisition was similar for both systems (59% and 56% improvement, compared with baseline, respectively). Despite an early performance decrement (45% for VR and 17% for VT at 13.2 ± 11.8 days post training) the acquired skill persisted over a mean 7-month follow-up period. While the early skill loss may seem substantial, little is known regarding skill retention after simulator-based surgical training. Grober et al\textsuperscript{25} showed that retention of skills was better at 4 months after hands-on bench model training for rodent vas deferens anastamoses, compared with didactic training alone. Torkington et al\textsuperscript{26} showed deterioration in basic laparoscopic skill according to MIST-VR testing using the manipulate diathermy task (25% over 3 months); however, training was conducted on a box trainer, not on the VR system. Anastakis et al\textsuperscript{27} found that a single session of a surgical skills curriculum did not confer long-term benefit to participants. Nonetheless, many more studies from nonsurgical fields have been published on durability of skill.\textsuperscript{15,17,18,21-23} Factors considered important for skill retention include the duration of retention interval, the quality and quantity of the original training, certain task characteristics, and individual differences.\textsuperscript{15}

While a meta-analysis of skill retention found that performance decay increased with longer retention intervals and documented a 92% skill loss at 1 year,\textsuperscript{15} this finding appears to be more applicable to cognitive tasks because several authors have shown long-term persistence of psychomotor skills.\textsuperscript{22,23} Likewise, in our study the retention interval did not appear to be related to skill deterioration because we found that skill acquired during a proficiency-based training persisted for many months despite an early decrement in the first few weeks post training.

The quality and quantity of training are also important factors for skill retention.\textsuperscript{15} Even though the quantity of training (number of repetitions) did not appear to be related to skill retention, the quality of the training in this study (proficiency based) may be accountable for the longevity of the acquired skill; having to achieve expert-derived performance may confer skills that are resistant to decay. On the other hand, it is unrealistic to believe that the trainee’s ability to achieve the expert-derived performance goals in 2 consecutive attempts (as was the endpoint of training for our subjects) provides them with enough skill to perform consistently at an expert level. Indeed, we showed in this study that while expert performance, as measured by simulators, remains stable and consistent over time, trainees lost 20% to 45% of the skill they achieved during proficiency-based training. Along those lines, the performance achieved during the posttest actually may reflect true trainee ability more accurately, compared with the performance achieved at the end of intensive training. Thus, even though we used a proficiency-based curriculum, it may not have been rigorous enough or ideal for maximal skill retention; alternative protocols for initial training or ongoing practice may be needed to maintain expert level performance.

Individual differences like baseline ability can also affect learning, as higher skilled individuals acquire and retain more skill over time.\textsuperscript{15,17,19} However, in our study, skill retention was not
Influenced by resident level even though senior residents outperformed junior residents at baseline on both tasks (indicating construct validity; statistically significant only for VT). This lack of correlation could be a consequence of our proficiency-based curriculum: All residents trained to the same expert-derived level, and their performance at training completion was by far superior to their baseline. In this context, acquired skill (quality of training) appears to be the primary determinant of retention, and baseline skill (resident level) may be of less importance, especially for basic laparoscopic skills.

Another factor, motivation, also plays an important role in skill acquisition and may influence long-term retention. Our baseline questionnaire data revealed that all participants were highly motivated because 100% expected to benefit from training. Motivation during training, however, may be affected by the type of task. In psychologic terms, “natural tasks are generally retained better than artificial tasks” because trainees exhibit a higher level of interest in learning tasks that seem “natural”. A study by Hamilton et al showed a clear preference of trainees for VT over VR, citing better visualization and tactile feedback that made VT more realistic. In regards to the significantly higher skill decay for the VR system, the superior face validity (realism) of the VT task may have motivated the trainees to maximize their learning and accounted for the lower skill deterioration on that task. In addition, interface differences between the 2 systems may have contributed to the skill retention differences at posttest. Although the VR training has proven effective in enhancing operative performance and incorporates sophisticated metrics, the environment is artificial and relatively gamelike, compared with VT. While some of the “tricks” for obtaining a good score easily may be forgotten for both VR and VT (possibly explaining the early performance decay), the VR system seemed to have many more such tricks and was much less forgiving as judged by the higher skill loss at posttest (45% vs 17%, respectively). Nevertheless, performance, as measured by both simulators, remained stable over a long period of time, substantiating their value as training tools. This knowledge should be taken into consideration when simulators are used as training or testing tools. Undoubtedly, as technology continues to improve, so too will the fidelity of the simulators, especially for VR.

While the findings of this study may not be generalized easily to the clinical setting, the results have important implications for training curricula design for residents and continuing education for practicing surgeons. The knowledge of how much skill is retained and when deterioration occurs (if it does) allows educators to plan retraining sessions at appropriate time frames to maintain optimal performance. Moreover, applying automaticity theory in training (how experts become automated at performing a task) and defining proficiency in new multidimensional fashions may prove more effective. These issues may be of lesser importance when, after initial training, a surgeon has ongoing clinical practice on a task that closely resembles the simulator task (eg, laparoscopic suturing); in that instance the surgeon’s skill will be unlikely to deteriorate because of ongoing practice. However, skill decay occurs potentially when a surgeon neither performs a procedure nor uses a specific skill for a long time; refresher training on a simulator, such as just before the procedure, might be of great value. In addition, although simulator training has proven to improve operative performance, the longevity of that benefit is unknown. Such studies are currently underway in our laboratory.

CONCLUSION
This study clearly shows long-term benefit for trainees after proficiency-based training on simulators. Additional studies are warranted to further optimize curriculum design and enhance acquisition of durable skills.

REFERENCES


