REVIEW ARTICLE

MEDICAL EDUCATION

Malcolm Cox, M.D., and David M. Irby, Ph.D., Editors

Teaching Surgical Skills — Changes in the Wind

Richard K. Reznick, M.D., M.Ed., and Helen MacRae, M.D.

From the Department of Surgery, University of Toronto, University Health Network (R.K.R.); and the University of Toronto Surgical Skills Centre at Mount Sinai Hospital (H.M.) — both in Toronto. Address reprint requests to Dr. Reznick at the Department of Surgery, University of Toronto, Suite 311, 100 College St., Toronto, ON M5G 1L5, Canada, or at richard.reznick@utoronto.ca.

SIR WILLIAM HALSTED INTRODUCED A GERMAN-STYLE RESIDENCY TRAINing system with an emphasis on graded responsibility at Johns Hopkins Hospital in 1889.[1] This system remains the cornerstone of surgical training in North America more than a century later. However, advances in educational theory, as well as mounting pressures in the clinical environment, have led to questions about the reliance on this approach to teaching technical skills.

Those pressures include a move toward a shorter workweek for residents[2,3] and an emphasis on operating room efficiency, both of which diminish teaching time. Yet the patients in our teaching hospitals are generally much sicker and have more complex problems than in times past. The increasing complexity of cases and a greater emphasis on mitigating medical error limit a faculty's latitude in assisting residents with technical procedures.

Sheer volume of exposure, rather than specifically designed curricula, is the hallmark of current surgical training.[4] But as opportunities for learning through work with "real" patients have diminished, interest in laboratories with formal curricula, specifically designed to teach surgical skills, has increased dramatically. In this new model of surgical education, basic surgical skills are learned and practiced on models and simulators, with the aim of better preparing trainees for the operating room experience.[5-10]

These new training techniques are based on established theories of the ways in which motor skills are acquired and expertise is developed. Fitts and Posner's three-stage theory of motor skill acquisition is widely accepted in both the motor skills literature and the surgical literature (Table 1).[11,12] In the cognitive stage, the learner intellectualizes the task; performance is erratic, and the procedure is carried out in distinct steps. For example, with a surgical skill as simple as tying a knot, in the cognitive stage the learner must understand the mechanics of the skill — how to hold the tie, how to place the throws, and how to move the hands. With practice and feedback, the learner reaches the integrative stage, in which knowledge is translated into appropriate motor behavior. The learner is still thinking about how to move the hands and hold the tie but is able to execute the task more fluidly, with fewer interruptions. In the autonomous stage, practice gradually results in smooth performance. The learner no longer needs to think about how to execute this particular task and can concentrate on other aspects of the procedure.

This model has obvious implications for surgical training. The earlier stages of teaching technical skills should take place outside the operating room; practice is the rule until automaticity in basic skills is achieved. This mastery of basic skills allows trainees to focus on more complex issues, both technical and nontechnical, in the operating room. To return to the example of knot tying, the learner who still

Table 1. The Fitts–Posner Three-Stage Theory of Motor Skill Acquisition.*			
Stage	**Goal**	**Activity**	**Performance**
Cognition	Understand the task	Explanation, demonstration	Erratic, distinct steps
Integration	Comprehend and perform mechanics	Deliberate practice, feedback	More fluid, fewer interruptions
Automation	Perform the task with speed, efficiency, and precision	Automated performance requiring little cognitive input, focus on refining performance	Continuous, fluid, adaptive

* Adapted from Fitts and Posner.[11]

has to think about how to tie a square knot is much less likely to pick up on other teaching that transpires in the operating room than is the learner who has mastered this simple skill.

Ericsson has helped to elucidate the acquisition of expertise.[13,14] Expert performance represents the highest level of skill acquisition and the final result of a gradual improvement in performance through extended experience in a given domain. According to Ericsson, most professionals reach a stable, average level of performance and maintain this status for the rest of their careers. In surgery, "experts" have been defined by Ericsson as experienced surgeons with consistently better outcomes than nonexperts. An extensive literature on the relationship of operative volume to clinical outcomes supports the hypothesis that practice is an important determinant of outcome[15]; the literature also provides support for Ericsson's contention that many professionals probably do not attain true expertise. However, volume alone does not account for the skill level among practitioners, since variations in performance have been shown among surgeons with high and very high volumes. Deliberate practice is a critical process for the development of mastery or expertise. Ericsson argues that the number of hours spent in deliberate practice, rather than just hours spent in surgery, is an important determinant of the level of expertise.[13]

Deliberate practice calls for the individual to focus on a defined task, typically identified by a teacher, to improve particular aspects of performance; it involves repeated practice along with coaching and immediate feedback on performance. The attained level of expertise has been shown to be closely related to time devoted to deliberate practice in the performance of expert musicians, chess players, and athletes. In the current model of surgical training, based primarily on apprenticeship, the opportunities for deliberate practice are rare. Operations are complex, and it is difficult to focus on one small component of the procedure.

In our opinion, in order to better plan instruction and assess the efficacy of curricular interventions, valid and reliable assessments of technical skills are needed. Evaluating performance in the operating room is difficult,[16] and most efforts have focused on techniques that standardize the assessment process outside the operating room. One such method is the Objective Structured Assessment of Technical Skills (OSATS),[17,18] in which candidates perform a series of standardized surgical tasks on inanimate models under the direct observation of an expert. Examiners score candidates using two methods. The first is a task-specific checklist consisting of 10 to 30 specific surgical maneuvers that have been deemed essential elements of the procedure. The second is a global rating form, which includes five to eight surgical behaviors, such as respect for tissues, economy of motion, and appropriate use of assistants. The validity and reliability of the OSATS are similar to those of the more traditional Objective Structured Clinical Examination (OSCE) and are acceptable for summative high-stakes evaluation purposes.[19-21] To date, we have created more than 40 OSATS stations; some examples are shown in Figure 1.

Other methods of assessment include the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS)[22] and the Imperial College Surgical Assessment Device (ICSAD).[23,24] Developed at McGill University in Montreal, the MISTELS uses an inanimate box to simulate the generic skills needed in the performance of laparoscopic surgery. It has been shown to be a valid and reliable instrument for assessing laparoscopic skills.[22] The ICSAD, developed at Imperial College in London, tracks hand motion using sensors placed on the trainee's hands during

Figure 1. Examples of OSATS Stations.

Examinees rotate through multiple stations, where they perform elements of surgical tasks and are graded by expert examiners using global rating forms and task-specific checklists. These examples are drawn from an "inventory" of more than 40 such stations.

the performance of a task. The sensors translate movement into a computerized tracing of hand motion, which provides an effective index of technical skill in both laparoscopic[23] and open[24,25] procedures. This index has been shown to have good concordance with OSATS scores.

Most surgical training programs make use of a variety of models, including inanimate models, virtual reality, live animals, and human cadavers, to simulate living human tissue and anatomy, as well as high-performance patient simulators for critical-incident and team training. Although human cadavers most closely approximate reality, their cost and limited availability, as well as the poor compliance of cadaveric tissue, limit their use. The use of live animals is also problematic because of ethical concerns, high costs, and the need for specialized facilities. In contrast, inanimate models are safe, reproducible, portable, readily available, and generally more cost-effective than

animals or cadavers. Some of the advantages and disadvantages of various models are summarized in Table 2.

Recent advances in virtual reality technology have demonstrated its potential for enhancing surgical skills training, and many virtual reality systems are now commercially available. Virtual reality provides the opportunity for very detailed feedback and may allow for more subtle measurement of trainee performance than is possible in the real world.[25] Measures of precision and accuracy as well as error rates can be calculated easily.[23,25,26]

Two prospective trials have demonstrated that residents who have been trained on low-fidelity (not very lifelike) virtual reality models (laparoscopic box trainers) make fewer intraoperative errors when performing a laparoscopic cholecystectomy than do residents who have not had the benefit of simulation training.[27,28] High-fidelity (lifelike) virtual reality models are also avail-

Table 2. Types of Simulations Available.

Simulation	Advantages	Disadvantages	Best Use
Bench models	Cheap, portable, reusable, minimal risks	Acceptance by trainees; low fidelity; basic tasks, not operations	Basic skills for novice learners, discrete skills
Live animals	High fidelity, availability, can practice hemostasis and entire operations	Cost, special facilities and personnel required, ethical concerns, single use, anatomical differences	Advanced procedural knowledge, procedures in which blood flow is important, dissection skills
Cadavers	High fidelity, only "true" anatomy simulator currently, can practice entire operations	Cost, availability, single use, compliance of tissue, infection risk	Advanced procedural knowledge, dissection, continuing medical education
Human performance simulators	Reusable, high fidelity, data capture, interactivity	Cost, maintenance, and downtime; limited "technical" applications	Team training, crisis management
Virtual reality surgical simulators	Reusable, data capture, minimal setup time	Cost, maintenance, and downtime; acceptance by trainees; three dimensions not well simulated	Basic laparoscopic skills, endoscopic and transcutaneous procedural skills

able for training in procedures such as colonoscopy and carotid artery stenting.[29] A Food and Drug Administration panel recently recommended the use of virtual reality simulation as an integral component of a training package for carotid artery stenting.[30]

However, high-fidelity virtual reality comes at a price. As a general rule, the higher the fidelity and the more realistic the model, the more expensive the training tool. Further investigation will be required to determine whether such an investment is worthwhile. A large carotid artery stenting trial that is currently under way will help assess the effectiveness of high-fidelity virtual reality models in training experienced practitioners.[31]

Fidelity may be less important at relatively junior levels of training. For example, when one group of medical students was trained with the use of a high-fidelity-video endoscopic urology system and another with the use of a simple bench model, the two groups showed the same improvement in performance and showed more improvement than the control group (given didactic training).[32] Likewise, among first-year surgical residents, improvement in performance for a variety of open procedures has been shown to be the same whether low-fidelity bench models or cadavers are used,[33] and residents working with a simple Silastic tubing model performed similarly to those working with the vas deferens of a live rat.[34]

What about the overall effectiveness of ex vivo surgical skills training? To date, the evidence for transfer to the operating room is stronger for minimally invasive surgery than for more traditional open procedures. In a series of experiments involving more than 200 surgeons and trainees, practice on a physical laparoscopic simulator led to the acquisition of skills that were transferable to complex laparoscopic tasks such as suturing.[22] Similarly, second- and third-year residents who received formal training on a physical laparoscopic simulator had a significantly greater improvement in video-trainer scores and global assessments of performance of a laparoscopic cholecystectomy than did residents who had no simulator training.[35]

The transfer of skills learned on virtual reality laparoscopic simulators has also been encouraging. Residents who received virtual reality training performed the dissection more quickly, made fewer errors, and had higher economy-of-movement scores during a laparoscopic cholecystectomy than did residents without such training.[27,28] Taken together, these studies strongly suggest that ex vivo laparoscopic training leads to detectable benefits for learners in clinical settings.

However, it remains unclear whether the improvement in performance after ex vivo training is durable. One study, by Grober and colleagues,[36] showed a durable positive effect of bench-model training in the task of microvascular anastomosis. In contrast, Sedlack and Kolars[37] found an initial positive effect of virtual reality training in colonoscopy. However, in this high-volume unit, where fellows were performing approximately 15 colonoscopies per week, fellows who were trained

with virtual reality models performed similarly to those who were not after a threshold of 30 procedures. Nonetheless, one could argue that whereas the advantages of training on a simulator may be limited to early procedural experience, the enhanced early learning curve may allow educators to be more efficient with their time.

Virtual reality has the potential to enhance surgical-team training as well as technical skills training. In aviation, teamwork training with simulation has been instrumental in reducing errors.[38,39] The importance of teamwork in preventing medical error is well recognized, and simulator-based team training has been advocated as a possible preventive approach. Early research results have been promising. Simulator-based trauma-team training has been associated with enhanced performance.[40] Task completion and simulated survival rates have improved after emergency-team training on a simulator.[41] Simulator-based training for crisis management in anesthesia, both for individuals and for teams, has been shown to be effective,[42] and simulated operating room environments are being assessed for training in both technical skills and teamwork.[43]

In summary, the report card on simulation, while not definitively positive, does suggest that it is an important addition to the training arsenal. The effectiveness of simulation training has been demonstrated primarily for lower-level learners,[27,28,32-34] with a positive effect demonstrated for both laparoscopic[27,28,35] and open[33,34] procedures. Results to date have not been validated in large-scale studies, however, and additional research demonstrating the generalizability of findings across institutions is required. Additional research on simulation training for more senior learners, and for surgeons in practice, is needed. Further work is also needed to address important motor-learning issues, such as whether it is preferable to practice whole operations or to practice segments of operations and then build the whole from the segments, what practice schedules are optimal, and how to optimize the transfer of skills to the operating room. Simulator-based skills training is a relatively new area of research, and we are only beginning to build our knowledge in this domain.

As computing power expands and the cost of simulation equipment falls, it is likely that most, if not all, surgical training programs will be devoting substantial curricular time to simulator-based training. Increasing evidence of the efficacy of ex vivo training, coupled with societal pressure, will probably mean that future residents will need to demonstrate proficiency in basic techniques before being allowed to operate on patients.

With these and other changes in the wind, surgical educators will need to incorporate meaningful assessment into residency programs, using rigorous, reliable, and regular means of assessment for all relevant surgical skills. Tools with the requisite levels of reliability and validity for summative assessment include the OSATS and MISTELS programs. For example, the MISTELS program might be used to assess a resident's performance of basic laparoscopic skills. A requisite level of performance would be required before the resident would be allowed to perform a laparoscopic cholecystectomy. Similarly, OSATS stations could be used for open surgery. Residents would thus be trained in the laboratory until preset criteria had been met and would only then be allowed to participate in the graduated performance of procedures in patients. Competence-based advancement, rather than time served, would become standard in surgical training.

One of the challenges of a competence-based system of education and assessment that has received little attention is how to establish pass or fail standards for the performance of technical skills. This is a major deficit in the assessment literature that will require attention if a criterion-based system is implemented or if certification of technical ability is required before licensure. Given the advances in technology and the accruing evidence of their effectiveness, now is the time to take stock of the changes we can and must make to improve the assessment and training of surgeons in the future.

Videos showing training models for microvascular and vascular anastomosis are available with the full text of this article at www.nejm.org.

REFERENCES

1. Carter BN. The fruition of Halsted's concept of surgical training. Surgery 1952; 32:518-27.

2. Council Directive 93/104/EC. Official Journal of the European Communities 1993;L307:18-24.

3. Leach DC. A model for GME: shifting from process to outcomes — a progress report from the Accreditation Council for Graduate Medical Education. Med Educ 2004;38:12-4.

4. Haluck RS, Krummel TM. Computers and virtual reality for surgical education in the 21st century. Arch Surg 2000;135: 786-92.

5. Scallon SE, Fairholm DJ, Cochrane DD, Taylor DC. Evaluation of the operating room as a surgical teaching venue. Can J Surg 1992;35:173-6.

6. Hutchison C, Hamstra S, Leadbetter W. The University of Toronto Surgical Skills Centre opens. Focus Surg Educ 1998; 16:22-4.

7. Heppell J, Beauchamp G, Chollet A. Ten-year experience with a basic technical skills and perioperative management workshop for first-year residents. Can J Surg 1995;38:27-32.

8. Lossing AG, Hatswell EM, Gilas T, Reznick RK, Smith LC. A technical-skills course for 1st-year residents in general surgery: a descriptive study. Can J Surg 1992;35:536-40.

9. Cauraugh JH, Martin M, Martin KK. Modeling surgical expertise for motor skill acquisition. Am J Surg 1999;177:331-6.

10. Reznick RK. Teaching and testing technical skills. Am J Surg 1993;165:358-61.

11. Fitts PM, Posner MI. Human performance. Belmont, CA: Brooks/Cole, 1967.

12. Kopta JA. The development of motor skills in orthopaedic education. Clin Orthop 1971;75:80-5.

13. Ericsson KA. The acquisition of expert performance: an introduction to some of the issues. In: Ericsson KA, ed. The road to excellence: the acquisition of expert performance in the arts and sciences, sports, and games. Mahwah, NJ: Lawrence Erlbaum Associates, 1996:1-50.

14. Idem. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79:Suppl 10:S70-S81.

15. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? A systematic review and methodologic critique of the literature. Ann Intern Med 2002;137:511-20.

16. Winckel CP, Reznick RK, Cohen R, Taylor B. Reliability and construct validity of a structured technical skills assessment form. Am J Surg 1994;167:423-7.

17. Martin JA, Regehr G, Reznick R, et al. Objective Structured Assessment of Technical Skill (OSATS) for surgical residents. Br J Surg 1997;84:273-8.

18. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. Am J Surg 1997;173:226-30.

19. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. Acad Med 1998;73:993-7.

20. Jansen JJ, Tan LH, van der Vleuten CP, van Luijk SJ, Rethans JJ, Grol RP. Assessment of competence in technical clinical skills of general practitioners. Med Educ 1995;29:247-53.

21. Szalay D, MacRae H, Regehr G, Reznick R. Using operative outcome to assess technical skill. Am J Surg 2000;180:234-7.

22. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. Ann Surg 2004;240: 518-28.

23. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator. Stud Health Technol Inform 1998;50:124-30.

24. Datta V, Mackay SD, Mandalia M, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. J Am Coll Surg 2001;193:479-85.

25. Darzi A, Mackay S. Assessment of surgical competence. Qual Health Care 2001; 10:Suppl 2:ii64-ii69.

26. Risucci D, Cohen JA, Garbus JE, Goldstein M, Cohen MG. The effects of practice and instruction on speed and accuracy during resident acquisition of simulated laparoscopic skills. Curr Surg 2001;58:230-5.

27. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002;236:458-63.

28. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 2004;91:146-50.

29. Cotin S, Dawson SL, Meglan D, et al. ICTS, an interventional cardiology training system. In: Westwood JD, Hoffman HM, Mogel GT, Robb RA, Stredney D, eds. Medicine meets virtual reality 2000. Washington, DC: IOS Press, 2000:59-65.

30. Gallagher AG, Cates CU. Approval of virtual reality training for carotid stenting: what this means for procedural-based medicine. JAMA 2004;292:3024-6.

31. Yadav JS, Wholey MH, Kuntz RE, et al. Protected carotid-artery stenting versus endarterectomy in high-risk patients. N Engl J Med 2004;351:1493-500.

32. Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. The effect of bench model fidelity on endourological skills: a randomized controlled study. J Urol 2002; 167:1243-7.

33. Anastakis DJ, Regehr G, Reznick RK, et al. Assessment of technical skills transfer from the bench training model to the human model. Am J Surg 1999;177:167-70.

34. Grober ED, Hamstra SJ, Wanzel KR, et al. The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures. Ann Surg 2004;240: 374-81.

35. Scott DJ, Bergen PC, Rege RV, et al. Laparoscopic training on bench models: better and more cost effective than operating room experience? J Am Coll Surg 2000;191:272-83.

36. Grober ED, Hamstra SJ, Wanzel KR, et al. Laboratory based training in urologic microsurgery with bench model simulators: a randomized controlled trial evaluating the durability of technical skill. J Urol 2004;172:378-81.

37. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: results of a pilot study. Am J Gastroenterol 2004;99:33-7.

38. Helmreich RL. Managing human error in aviation. Sci Am 1997;276:62-7.

39. Leedom D, Dimon R. Improving team coordination: a case for behavior-based training. Mil Psych 1995;7:109-22.

40. Holcomb JB, Dumire RD, Crommett JW, et al. Evaluation of trauma team performance using an advanced human patient simulator for resuscitation training. J Trauma 2002;52:1078-85.

41. DeVita MA, Schaefer J, Lutz J, Wang H, Dongilli T. Improving medical emergency team (MET) performance using a novel curriculum and a computerized human patient simulator. Qual Saf Health Care 2005; 14:326-31.

42. Wong AK. Full scale computer simulators in anesthesia training and evaluation. Can J Anaesth 2004;51:455-64.

43. Moorthy K, Munz Y, Adams S, Pandey V, Darzi A. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. Ann Surg 2005;242:631-9.

Copyright © 2006 Massachusetts Medical Society.