Teaching cognitive skills improves learning in surgical skills courses: a blinded, prospective, randomized study

Julie A. Kohls-Gatzoulis, MD;[*†] Glenn Regehr, PhD;[*] Carol Hutchison, MD[*‡]

Objective: To investigate the teaching of cognitive skills within a technical skills course, we carried out a blinded, randomized prospective study. **Methods:** Twenty-one junior residents (postgraduate years 1–3) from a single program at a surgical-skills training centre were randomized to 2 surgical skills courses teaching total knee arthroplasty. One course taught only technical skill and had more repetitions of the task (5 or 6). The other focused more on developing cognitive skills and had fewer task repetitions (3 or 4). All were tested with the Objective Structured Assessment of Technical Skill (OSATS) both before and after the course, as well as a pre- and postcourse error-detection exam and a postcourse exam with multiple-choice questions (MCQs) to test their cognitive skills. **Results:** Both groups' technical skills as assessed by OSATS were equivalent, both pre- and postcourse. Taking their courses improved the technical skills of both groups (OSATS, $p < 0.01$) over their pre-course scores. Both groups demonstrated equivalent levels of knowledge on the MCQ exam, but the cognitive group scored better on the error-detection test ($p = 0.02$). **Conclusions:** Cognitive skills training enhances the ability to correctly execute a surgical skill. Furthermore, specific training and practice are required to develop procedural knowledge into appropriate cognitive skills. Surgeons need to be trained to judge the correctness of their actions.

Objective : Nous avons procédé à une étude prospective randomisée à l'insu pour étudier l'enseignement des compétences cognitives dans le contexte d'un cours de compétences techniques. **Méthodes :** On a réparti au hasard 21 résidents débutants (années de formation postdoctorale 1 à 3) d'un même programme offert dans un centre de formation en techniques chirurgicales entre deux cours de formation en techniques chirurgicales où l'on enseignait l'arthroplastie totale du genou. Dans l'un des cours, on enseignait la technique seulement et le cours prévoyait plus de répétitions de la tâche (5 ou 6). Dans l'autre cours, on insistait davantage sur le développement de compétences cognitives, et les répétitions étaient moins nombreuses (3 ou 4). Tous se sont soumis au test d'évaluation structurée objective des compétences techniques (OSATS) avant et après le cours, ainsi qu'à un examen de détection des erreurs avant et après le cours et à un examen consécutif au cours comportant des questions à choix multiples (QCM) visant à vérifier leurs compétences cognitives. **Résultats :** Les compétences techniques des deux groupes évalués par le test OSATS s'équivalaient, à la fois avant et après le cours. Le cours a amélioré les compétences techniques dans les deux groupes (OSATS, $p < 0,01$) par rapport aux résultats antérieurs. L'examen QCM a démontré que les deux groupes possédaient des connaissances équivalentes, mais le groupe axé sur les compétences cognitives a obtenu de meilleurs résultats au test de détection des erreurs ($p = 0,02$). **Conclusions :** La formation en compétences cognitives améliore la capacité d'exécuter correctement une technique chirurgicale. De plus, une formation spécifique et la pratique s'imposent pour intégrer la connaissance de l'intervention aux compétences cognitives appropriées. Il faut donner aux chirurgiens la formation dont ils ont besoin pour savoir juger que leurs interventions sont correctes.

*From the *Centre for Research in Education, University Health Network, University of Toronto, Toronto, Ont., †Royal National Orthopædic Hospital Trust, Stanmore, England, and ‡San Ramon Regional Medical Center, San Leandro, Calif.*

This work has been presented at the American Academy of Orthopædic Surgeons Annual Meeting, Orlando, Fla., March 15, 2000; Association for Surgical Education Annual Meeting, Toronto, Ont., May 4–6, 2000; and Canadian Orthopædic Association Annual Meeting, Edmonton, Alta., June 4–6, 2000.

Accepted for publication Mar. 24, 2004

Correspondence to: *Dr. Julie A. Kohls-Gatzoulis, gatzoulis@ntlworld.com*

An emerging consensus in the surgical specialties is that technical skill acquisition should have a greater emphasis during surgical training.[1–5] The central theme is that surgical training programs and academies should provide formal structured forums for those learning new procedures to practise their technical skills before using them in the intense environment of the operating room. In response, many laboratory-based skills-development courses have been instituted.[6–12] This focus on the technical aspect of surgery is quite appropriate. Research has suggested that time spent on tasks is an important factor in attaining motor-skill proficiency.[13,14]

Although a central purpose of surgical skills courses is practice in a safe environment, paradoxically, many courses continue to have a heavy focus on didactic teaching. The lectures may describe the procedural steps, but the actual periods scheduled for hands-on practice can be very limited and tend to depend on the availability of time and resources rather than on the needs of course attendees.

Recognizing the importance of practice, some surgical educators are beginning to create courses that allow for both initial didactic teaching and large blocks of practice time. However, the success of a surgical procedure depends on more than the surgeon's ability to perform each of the manœuvres associated with the procedure: Cognitive skills such as error detection, forward planning and decision-making are also crucial. The addition of practice and feedback in the cognitive domain of a procedure would necessarily result in some loss of opportunity for practice and feedback in its technical skills. Finding the balance between these aspects of training might not be particularly easy.

This study was an attempt to understand the role of cognitive skills in skills-oriented courses. Its primary purpose was to determine whether a surgical skills course that integrates cognitive and technical skills training produces greater combined skill than one that is purely technical. In adding cognitive skills training to a surgical skills course, it was recognized that there would be less time for technical skills practice, in comparison with a course of the same duration focusing only on technical skills. Also, the possibility existed that this dual focus would overload trainees, and actually reduce overall skill in those individuals who had practised both technical and cognitive skills. Because of the number of questions raised, we felt a randomized prospective study comparing technical skills practice with combined cognitive and technical skills practice would be necessary in order to draw meaningful conclusions.

Although an established measure of technical skill, the Objective Structured Assessment of Technical Skill[15–18] (OSATS), exists, no evaluation tool has been available to study cognitive skills related to surgery. Thus, a secondary implication of this project was the introduction of a new measure designed to evaluate overall cognitive skills. The new measure tested skillfulness at detecting common surgical errors related to the procedure (presenting examples of outcomes from poorly performed procedures and evaluating residents' ability to identify the error, its implications and the method for correcting it).

We selected total knee arthroplasty (TKA) as the topic for our skills courses because of the senior investigator's area of subspecialization (hip and knee arthroplasty). Also, we had identified trainee interest in the topic owing to the complexity of this nevertheless common orthopedic procedure.

Methods

Ethical approval for the study was obtained from the human-subjects review committee at our institution and university.

Study subjects

Sample size for the study was determined with a power calculation. We had no previous data to help us predict the difference that we would observe; however, a large difference between the 2 groups was expected on the cognitive measures. Using an α value of 0.05, a β value of 0.2, and a δ value of 1.5 standard deviations, the power calculation yielded a group size of 12 residents.

The 30 junior residents enrolled in the Orthopedic Surgery Program at our institution were sent a letter explaining the study along with a consent form. To be eligible for the study, they were required to be available to attend the entire (weekend) course. Subjects who volunteered and could meet this criterion were stratified by year of training and randomized into 2 study groups.

Study personnel

Eighteen consultant orthopedic surgeons volunteered to act as examiners for the pre-course test and the final evaluation of technical skills (each 3.5 h long). Six examiners came to both tests; 6 came to the pre-course test only; and 6 to the postcourse test only.

The exam assistants recruited were trained operating-room nurses and operating-room assistants, paid (at hospital rates) for 3.5 hours at each iteration of the test. They were given an orientation to the instrumentation before the examination.

Six orthopedic surgeons, specializing in arthroplasty and known to be especially good at teaching residents, were recruited as instructors. Instructors donated their time to teach the 14-hour courses; they were randomized before the course and given an orientation on the teaching method to be used for the course they were instructing. After the course, the instructors were sent feedback from teaching evaluations completed by course attendees.

Educational materials

The large amount of material required for the courses was supplied to instructors a week beforehand. To prepare faculty development handouts, a literature search was done on structuring practice and feedback for motor learning. A video on the use of the Zimmer NexGen knee system was supplied by Zimmer, Canada (Mississauga, Ont.). A professionally formatted handout on TKA was prepared containing information on the clinical reasoning needed to decide when a TKA is indicated, how to do the procedure, and its possible complications. Slides for the introductory TKA lecture given to both groups were prepared on Powerpoint, with handouts for residents so that they could follow along. Overheads were made to help guide instructors in the lecture on error detection.

Outcome measures

Technical skills

The technical skills test used a pre-/postcourse design to establish the amount of technical improvement obtained from each of the 2 teaching approaches. The technical test followed the OSATS format.[15–17] Residents' technical skill was evaluated by qualified orthopedic surgeons using a task-specific checklist, a global rating scale and an end-product analysis.[18] On the pre-course test, each trainee performed a TKA on 2 "normal" Sawbones knee models (Pacific Research Laboratories, Vashon, Wash.); all teaching and testing before the postcourse test was done on models of normal knees. The postcourse test included a TKA on 1 normal model and 1 arthritic model.

Knowledge and cognitive skills

Two other evaluations were done after the course only, because of concern that pre-course evaluation of these skills might contaminate the residents' learning by encouraging them to seek out additional information. These 2 measures were an exam with 40 multiple-choice questions (MCQs), designed to test knowledge related to the performance and preoperative planning of TKA; and a new 40-minute, 8-station error-detection test developed to evaluate the degree to which the residents could identify errors in surgery, the implications of those errors for patient outcome and the best method for fixing them once they occur.

Models for the error-detection test prepared before the study each contained a single obvious error of a magnitude larger than what would hopefully actually ever occur. Examples include an anterior (rather than a posterior) slope on the tibia, a tibia cut of 20° varus, and internal rotation of the femoral component by 20°. The models were tested on experienced arthroplasty surgeons and fellows; the surgeons scored perfectly, and the fellows near-perfectly. Although this was only a first step toward a demonstration of validity, we thought it sufficient for use of this error-detection test in the context of our experiment. The test was marked independently by 2 orthopedic examiners blinded to group and course attendance.

Procedure

The study ran over a single weekend. On the Friday afternoon, both groups took the pre-course test, which consisted of 2 stations. At each station was a model knee on which the trainee performed a TKA. An examiner was present to evaluate the trainee's technical performance throughout the procedure. The exam assistant was instructed to provide only the aid that was requested.

The following morning, a 1-hour lecture on TKA was given jointly to both groups by an experienced arthroplasty surgeon. The areas covered included not only patient selection and the steps involved in the procedure, but also surgical pitfalls and errors and possible surgical complications. All subjects started in the skills lab as 1 large group and received instruction about the instrumentation. All subjects also received a technical demonstration by another experienced arthroplasty surgeon on using the instrumentation to do a TKA on a model knee.

At this point, the subjects were separated into the 2 groups. The technical group was able to repeat the procedure more times (5 or 6 times) and therefore experienced less encouragement to apply cognitive skills. The cognitive group spent more time assessing their finished product, at the expense of the number of times they did the procedure (only 3–4 times).

In both groups, all residents were required at the completion of each procedure to assess the quality of their completed knee and then to have an instructor do the same. The evaluations were kept in a booklet given to each trainee at the start of the course, to encourage self-assessment among the residents, to provide feedback from the instructors and to keep track of the number of TKAs performed in the 2 groups. The technical group evaluated their end product on its technical quality, e.g., the precision with which the bony cuts were made. The cognitive group evaluated their end product on its overall quality, e.g., the accuracy with which the bony cuts were placed.

Instructors for the technical group were directed to show the residents how to improve their technical skills. Examples of this would be demonstrating how to keep the saw blade from flexing in the jig, or the placement of the femoral component in 3° of external rotation (without emphasizing how this improves patellar tracking, which was covered in the common introductory lecture). However, if a trainee raised a question related to cognitive skills or decision-making, it was answered fully by the instructor. We did this because we recognized that the goal of either course was to improve surgical skills, and that even as we tried artificially to prevent such learning, trainees who are resourceful enough may find answers from sources other than instructors.

Instructors for the cognitive group were directed to encourage residents to examine their finished product for errors such as an internally rotated component or varus joint line. As well, the cognitive group had an additional 1-hour lecture that covered again the material presented in the initial lecture given to both groups demonstrating surgical errors related to TKA and re-emphasizing how these surgical errors produce clinical problems (internal rotation of the femoral component, a tight flexion gap, etc.).

On the Sunday morning, both groups received a refresher video demonstration on how to do the procedure before breaking into their respective groups to continue practising.

On the Sunday afternoon, sets of postcourse technical and cognitive skills tests were given to both groups. To ensure that repeat OSATS examiners were not biased by the pre-course test, we placed them in different exam tracks so that they did not examine the same residents. The residents expected that they would take the postcourse OSATS, but the postcourse multiple-choice and error-detection tests were unexpected.

Models for the error-detection test contained errors of such a grave nature as to be incompatible with a successful surgical outcome. Information about how to prevent these errors, like the information being tested on the MCQ test, was covered in the common introductory lecture. Trainees had also had the chance to examine their own knee replacements for errors during the course.

Statistical analysis

For measures of knowledge and cognitive skills, an independent *t* test (2-tailed) was used to compare the 2 groups. On the technical OSATS, 2-tailed *t* tests were used to compare the 2 groups for the pre-course test (test results for the 2 TKAs were combined) and the postcourse test (results for the TKAs performed on the normal and the arthritic models were kept separate).

Results

Of the 30 eligible junior residents enrolled in the Orthopædic Surgery Program at our institution, 26 originally volunteered for the study. The 4 residents who did not sign up had previous commitments that prevented them from attending the study. Of those who signed up, 5 residents had on-call scheduling difficulties and pulled out of the study at the last minute. The final number of residents participating in the study was 21, all from the first 3 years of training. (We considered inviting more senior trainees from our institution or inviting trainees from a nearby training program to attend the courses, but decided that having trainees from a different training background participate in the research could create further difficulties.)

Group composition for the technical group was 3 postgraduate year 1s, 6 year 2s, and 2 year 3s; for the cognitive group, 3 postgraduate year 1s, 5 year 1s, and 2 year 3s. All residents attended the entire course and the testing.

Residents evaluated both of the courses highly. Despite having performed better, some feedback from trainees in the cognitive group indicated a desire to repeat the procedure more times. Trainees from the technical group were uniformly happy with the course they received. Participants were informed of study means in addition to their own individual results.

Six instructors agreed to be present for the day-and-a-half that the course ran. Those teaching the technical group included 2 academic arthroplasty surgeons and 1 newly accredited surgeon training in arthroplasty. Those instructing the cognitive group included 1 academic and 1 community arthroplasty surgeon, and 1 newly accredited surgeon training in arthroplasty. Because of poor weather, the community arthroplasty surgeon was unable to make it to the course, leaving only 2 surgeons to teach the cognitive group. On the second morning, the academic arthroplasty surgeon did not attend, leaving only 1 instructor for the cognitive group. One of the academic surgeons who had taught the technical group the previous day was asked to switch over. The instructors were highly rated by all residents at the end of the course.

Table 1 summarizes results for the tests of technical skill. On both pre-course and postcourse tests, differences between the 2 groups on any of the measures of technical skill (checklist, global rating scale or the end-product analysis) were not significant.

Recognizing that the lack of statistical significance might have resulted from an underpowering of our study, we ran power calculations for each of the comparisons to determine the number of subjects per group that would be required for significance, given the size of the differences seen. Assuming an α of 0.05, we determined that the number of subjects for each group required to achieve statistical significance with an 80% power ranged from 92 per group (for the checklist score on the arthritic knee) to 2719 per group (for the checklist score on the normal knee). Thus, we have some confidence in assuming that the failure to see a significant difference between the 2 groups was more a function of the small, clinically nonsignificant differences seen than a function of an underpowering of our study.

In contrast, the difference between the performance of both groups from pre-course to postcourse test was significant for all measures: checklist *t*

= 11.4, $p < 0.01$; global rating scale $t = 10.9$, $p < 0.01$; end-product analysis $t = 10.2$, $p < 0.01$. In addition, although there was a trend toward slightly lower scores on the arthritic model, there was no statistically significant difference in the postcourse tests between the normal and arthritic models on any of the measures used.

Table 1 also shows the data for the 2 groups on the measure of knowledge: the multiple-choice exam. No significant difference was found between the 2 groups on this measure. On the error-detection test, however, the cognitive group performed significantly better than the technical group: $t = 2.67$, $p = 0.02$. Interrater reliability between the 2 examiners who marked the error-detection test was 0.92. It was felt that the results from the experienced arthroplasty surgeon were more likely to reflect the true score (results in Table 1 are those of the experienced examiner).

Discussion

Our purpose was to delineate factors that could help guide surgical educators in the structuring and design of technical skills courses. Substantial information comes from this study. The discussion will be structured as follows: first the implications for technical skill acquisition will be discussed, then general knowledge and cognitive skills acquisition, followed by the effects of combining technical and cognitive skills acquisition. Finally, future directions will also be considered.

The pre-course test–postcourse test design of our study is very beneficial for educators working in all surgical specialties, especially when one considers that many Canadian programs are directing large amounts of funding toward the establishment of surgical skills centres, on the assumption that training in such an environment will indeed lead to improved acquisition of technical skills. In this study, the improvement from pre-course to postcourse OSATS scores was substantial ($p < 0.01$); this result supports the use of skills centres to teach complex surgical procedures to junior trainees. It is extremely important to emphasize the ample time allocated for practice in this study. Although the skills taught in our study were orthopedic skills, it is reasonable to extrapolate that complex procedures from other specialties can also be learned in a lab-based setting using a similar course design.

A question that often plagues surgical skills courses is whether the skills learned will be transferable to operating room. Transfer of technical skills is always difficult to address. Our use of a diseased model on the postcourse OSATS was meant to roughly address transfer. When presented with the unfamiliar model, many trainees did show a slight decrease in their OSATS global rating score (compared with normal-model postcourse OSATS), but this difference was not substantial enough to be statistically significant. We can conclude that the skill set learned could be generalized enough to transfer to a different, more complex model. We acknowledge that the question of how much transfer of technical skill there will be to the operating room remains unanswered, but in some ways that question is less relevant in this study. Because the skills being taught were of an advanced nature, the goal of teaching TKA to junior residents was to allow them to gain confidence and learn safe technique so that supervised training could safely begin in the operating room, rather than expecting the junior residents to completely transfer to performing TKAs on their own in the operating room.

The MCQ exam was meant to test knowledge at the end of the course. The questions were designed to test

Table 1

Scores on examinations for technical and cognitive skills

Test		Mean group score, % (& SD) Technical	Cognitive	Group difference,* % (& SD)
Pre-course technical skills (average of 2 normal knees)				
	Task-specific checklist	27 (24)	26 (26)	–1 (25)
	Global rating scale	22 (16)	19 (19)	–3 (17)
	End-product analysis	26 (32)	15 (15)	–11 (25)
Postcourse technical skills				
Normal knee	Task-specific checklist	83 (10)	83 (13)	0 (12)
	Global rating scale	70 (24)	72 (14)	2 (20)
	End-product analysis	81 (15)	80 (17)	–1 (16)
Arthritic knee	Task-specific checklist	88 (12)	77 (14)	–11 (13)
	Global rating scale	68 (19)	62 (17)	–6 (18)
	End-product analysis	83 (10)	78 (25)	–5 (19)
Difference from pre-course test				
Normal knee	Task-specific checklist	56 (25)	57 (28)	1 (26)
	Global rating scale	48 (24)	53 (19)	5 (22)
	End-product analysis	55 (40)	65 (26)	10 (34)
Arthritic knee	Task-specific checklist	61 (26)	51 (21)	–10 (24)
	Global rating scale	46 (21)	43 (24)	–3 (22)
	End-product analysis	57 (27)	63 (24)	8 (26)
Cognitive skills (post-course only)				
Multiple-choice		57 (16)	62 (10)	5 (13)
Error-detection		57 (20)	76 (10)	19 (16)†

*None of the differences seen between groups were statistically significant, except for †$p = 0.02$.
SD = standard deviation

an individual's understanding of the cognitive skills related to the procedure. Statistically, both groups scored the same; it is therefore highly likely that both groups had the same level of knowledge about the cognitive aspects of the procedure. This is interesting when the MCQ exam results are compared with the error-detection test results. The latter test required residents to not only know what is required for a successful outcome but to be able to recognize an example of a very poorly executed procedure when presented with one. The cognitive group scored significantly higher ($p = 0.02$) on the error-detection test. To summarize, both groups had similar knowledge levels related to the requirements of a successful procedure (as tested by the MCQ exam), but only the cognitive group had developed through practice the cognitive skills to actually recognize drastic examples of poorly performed THAs.

The work of Ericsson and colleagues[13,14] had led us to expect that cognitive skills, like technical skills, required practice to develop. The results do show just that. The usefulness of our study to surgical educators is that it shows that the setting of the surgical skills centre can be used to develop skills in the cognitive domain, as we had theorized. What is also useful is that the interaction between cognitive and technical skills acquisition has been unstudied. By comparing 2 otherwise equal groups, we have been able to show that the additional focus on cognitive skills was not harmful to the acquisition of technical skills (in the cognitive group). We feel that surgical educators from other surgical specialties can similarly expect to be able to develop cognitive skills through practice without adversely affecting the acquisition of technical skills during the same practice time.

The difference in the number of repetitions raises interesting questions. The cognitive group performed the procedure fewer times (3 or 4 times) than the technical group (5 or 6 times) and, perhaps unexpectedly, scored equivalently on the postcourse OSATS. Ericsson's and colleagues work[13,14] and also that of Cauraugh[19] suggest that efforts aimed at cognitive skills training might accelerate the understanding and planning of the procedure, giving meaning to the actions being practised, and reduce the overall training time necessary to become competent both cognitively and technically. Our results may support the work of Ericsson's group, but there are other possible explanations related to the groups' learning curves and to the sensitivity of OSATS as a measure of technical skill. Firstly, we do not know what the learning curves were for the 2 groups. It may be that the technical group had reached a plateau after doing the procedure 3–4 times. Conversely, it may be that the technical group was learning at a slower rate and required 5–6 repetitions to acquire the same level of skill that the cognitive group had acquired with fewer repetitions. Secondly, the sensitivity of OSATS as a measure of technical skill is unknown; it may be that there was a difference in the technical skill between the 2 groups that OSATS was unable to detect.

This discussion on measures of surgical skill leads to future directions of research. There are many people interested in being able to quantify surgical skill. At the end of our trial, we had 2 very different groups. Although both may have greatly improved their levels of technical skill, only the cognitive group had developed an ability to identify serious mistakes related to the procedure. Yet, OSATS could not detect the difference between these groups. The implication is that surgical educators must seek to create and standardize measures that examine more than technical skill. Had we not created a new measure, the error-detection test, we would have concluded that the 2 skills courses were equal, given that on every other measure the 2 groups were statistically equivalent. Clearly, further work to standardize and validate error-detection testing is worth exploration.

Cost is always a consideration, and this study shows that teaching cognitive skills is cost-effective. The cognitive group used fewer models. It should be emphasized that the cognitive group still did practice the procedure 3–4 times and also received structured feedback. We do not intend to reinforce an attitude of acceptance toward courses that do not allocate enough practice time, models or instructors. Perhaps it is interesting to note that some residents in the cognitive group reported feeling disadvantaged because they had performed fewer knee replacements than their counterparts in the technical group. Further research is needed, but we would advise educators to push trainees to gain all they can from each model used.

The most important implication of our research is the important benefit that junior residents received by the introduction of a cognitive-skills training component to skills-development courses. One could reasonably postulate that as well as total knee arthroplasty, many other surgical procedures would also require training and practice to produce cognitive skills and the ability to detect errors. Given the current societal pressure to reduce errors,[20,21] this result must be highlighted. To help prevent errors we must specifically teach cognitive skills and allow residents to practice detecting errors.

If we aim to maximize the effectiveness of skills-development courses for generating competence in a surgical procedure, we would do well to apply the same concepts of teaching, practice and feedback to the cognitive skills related to the procedure. The addition of cognitive skills training creates more effective learning and makes more efficient use of resources. Surgeons require specific training in order to judge the correctness of their actions.

Acknowledgements: This research was in part funded by a grant from the Association for Surgical Education of US$10 000. We would like to express our appreciation of the orthopedic surgeons who donated their time to participate as examiners and course instructors in this study. We would also like to thank Tom Dorman and Zimmer Canada for the loan of NexGen knee-system equipment trays.

Competing interests: None declared.

References

1. DesCôteaux JG, Leclère H. Learning surgical technical skills. *Can J Surg* 1994;38: 33-8.
2. Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993;165:358-63.
3. Barnes RW. Halstedian technique revisited. *Ann Surg* 1989;210:118-21.
4. Spencer FC. Teaching and measuring surgical techniques — the technical evaluation of competence. *Bull Am Coll of Surg* 1978;63:9-14.
5. Kopta JA. An approach to the evaluation of operative skills. *Surgery* 1971;70:297-303.
6. Cundiff GW. Analysis of the effectiveness of an endoscopy education program in improving resident's laparoscopic skills. *Obstet Gynecol* 1997;90:854-9.
7. Heppell J, Beauchamp G, Chollet A. 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. Steel RJC, Walder C, Herbert M. Psychomotor testing and the ability to perform an anastomosis in junior surgical trainees. *Br J Surg* 1992;79:1065-7.
10. Barnes RW. Surgical handicraft: Teaching and learning surgical skills. *Am J Surg* 1987;153:422-7.
11. Lazaro EJ, Rush BF, Blackwood JM, Swaminthan AP. Assessment of students' experiences in technical procedures in surgical clerkship. *Can J Surg* 1978;21:540-1.
12. Kopta JA. The development of motor skills in orthopædic education. *Clin Orthop* 1971;75:80-5.
13. Ericsson KA, editor. *The road to excellence: the acquisition of expert performance in the arts and sciences, sports and games.* Mahwah, NJ: Lawrence Erlbaum Associates; 1996.
14. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993;100:363-406.
15. Martin JA, Regehr G, Reznick RK, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84:273-8.
16. Reznick RK, Regehr G, MacRae, H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. *Am J Surg* 1997;173:226-30.
17. Faulkner H, Regehr G, Martin J, Reznick RK. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med* 1996;71:1363-5.
18. Regehr G, Macrae H, Reznick RK, Szalay D. Comparing the pschometric properties of checklists and global rating scales for assessment of performance on an OSCE-format examination. *Acad Med* 1998;73: 993-7.
19. Cauraugh JH. Modeling surgical expertise for motor skill acquisition. *Am J Surg* 1999;177:331-6.
20. Reason J. Human error: models and management. *BMJ* 2000;320:768-70.
21. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross-sectional surveys. *BMJ* 2000;320:745-9.