Cognitive task analysis for teaching technical skills in an inanimate surgical skills laboratory

George C. Velmahos, M.D. a,d,* Konstantinos G. Toutouzas, M.D. a, Lelan F. Sillin, M.D. a, Linda Chan, Ph.D. b, Richard E. Clark, Ed.D. c, Demetrios Theodorou, M.D. a, Fredric Maupin, Ed.D. c

aDepartment of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
bDepartment of Biostatistics, Los Angeles County and University of Southern California Medical Center, Los Angeles, CA, USA
cSchool of Education, University of Southern California, Los Angeles, CA, USA
dLos Angeles County and University of Southern California Medical Center, 1200 N. State St., Rm. 9900, Los Angeles, CA 90033, USA

Manuscript received September 3, 2002; revised manuscript December 24, 2002

Presented at the Annual Meeting of the Association for Surgical Education, Baltimore, Maryland, April 4–6, 2002

Abstract

Background: The teaching of surgical skills is based mostly on the traditional “see one, do one, teach one” resident-to-resident method. Surgical skills laboratories provide a new environment for teaching skills but their effectiveness has not been adequately tested. Cognitive task analysis is an innovative method to teach skills, used successfully in nonmedical fields. The objective of this study is to evaluate the effectiveness of a 3-hour surgical skills laboratory course on central venous catheterization (CVC), taught by the principles of cognitive task analysis to surgical interns.

Methods: Upon arrival to the Department of Surgery, 26 new interns were randomized to either receive a surgical skills laboratory course on CVC (“course” group, n = 12) or not (“traditional” group, n = 14). The course consisted mostly of hands-on training on inanimate CVC models. All interns took a 15-item multiple-choice question test on CVC at the beginning of the study. Within two and a half months all interns performed CVC on critically ill patients. The outcome measures were cognitive knowledge and technical-skill competence on CVC. These outcomes were assessed by a 14-item checklist evaluating the interns while performing CVC on a patient and by the 15-item multiple-choice-question test, which was repeated at that time.

Results: There were no differences between the two groups in the background characteristics of the interns or the patients having CVC. The scores at the initial multiple-choice test were similar (course: 7.33 ± 1.07, traditional: 8 ± 2.15, P = 0.944). However, the course interns scored significantly higher in the repeat test compared with the traditional interns (11 ± 1.86 versus 8.64 ± 1.82, P = 0.03). Also, the course interns achieved a higher score on the 14-item checklist (12.6 ± 1.1 versus 7.5 ± 2.2, P < 0.001). They required fewer attempts to find the vein (3.3 ± 2.2 versus 6.4 ± 4.2, P = 0.046) and showed a trend toward less time to complete the procedure (15.4 ± 9.5 versus 20.6 ± 9.1 minutes, P = 0.149).

Conclusions: A surgical skills laboratory course on CVC, taught by the principles of cognitive task analysis and using inanimate models, improves the knowledge and technical skills of new surgical interns on this task. © 2004 Excerpta Medica, Inc. All rights reserved.

Keywords: Skill; Cognition; Learning; Surgery; Model; Cognitive task analysis; Central vein; Catheterization

The methods to teach technical skills to surgical trainees have remained vague and underdeveloped [1]. In a survey of 152 program directors, approximately half of them felt that residents needed better training in the technical aspects of surgery [2]. For years, technical skills have been taught “in action” by the “see one, do one, teach one” method from resident to resident [3]. This teaching method is associated with great variability among teachers, inconsistent effectiveness, and potential risks for the patients used as subjects for training.

Teaching technical skills in a surgical skills laboratory has emerged in recent years as a new method of training
residents in a structured, stress-free environment [4–7]. Surgical skills laboratory curricula using artificial models or computerized simulation are increasingly gaining advantage over courses using animals, primarily because of cost issues and ethical considerations [1]. However, the methods of teaching technical skills in surgical skills laboratory are still inadequately defined. Furthermore, the effectiveness of such teaching is poorly documented and usually relies on subjective evaluations by the faculty or the trainees at the end of a course [3,5,8,9].

Cognitive task analysis is a relatively new educational method by which the automated elements of performing a procedure can be captured and taught effectively from experts to novices [10–12]. The method has been used with great success in private corporations and the military [12–14]. It has been recently tried in the health care field [15,16]. The objective of the current study is to examine the effectiveness of a surgical skills laboratory course on central venous catheterization (CVC) taught by the method of cognitive task analysis to new surgical interns. The effectiveness of the course is evaluated by assessing the competence of the trainees in performing CVC on patients under realistic clinical conditions.

Methods

Group allocation and course design

The 26 incoming preliminary and categorical interns of the Department of Surgery of the University of Southern California for the year 2001–2002 (including interns from other departments who spent an internship year at the Department of Surgery) were randomly assigned to two groups: a “traditional” group (n = 14) and a “course” group (n = 12). The interns of the first group were taught CVC by the traditional method: after observing the procedure, they practiced it on patients under the supervision of a more senior resident. The interns of the second group received a 3-hour course on CVC at our surgical skills laboratory. The course was given on the Saturday morning of June 30, 2001, and consisted of three parts: (1) for the first 15 minutes the interns were given a presentation on CVC, including indications, contraindications, anatomical considerations, technical points, and potential complications; (2) for the next 30 minutes CVC was demonstrated in detail by a senior resident (D.T.) on an artificial model (Laerdal IV Torso; Laerdal Medical Corp, Wappingers Falls, New York) according to cognitive task analysis principles; and (3) for the remaining 2 hours and 15 minutes the interns were allowed to practice on artificial models. Three stations were developed with one model and one instructor in each station. Four interns participated in each station, taking turns in practicing CVC of the internal jugular and subclavian veins. Each intern had the opportunity to insert a central venous catheter at least four times in the artificial model.

Cognitive task analysis

The teaching of CVC was done according to the principles of cognitive task analysis [12]. In preparation for the course, two experts on cognitive task analysis (R.E.C. and F. P.) met with an expert on CVC (G.C.V.) to analyze the cognitive elements of the task of CVC. Then, they met with a second expert on CVC (J.A.M.) and recorded differences between the two approaches. The differences were reconciled in a third meeting and a structured framework was developed for teaching CVC. A document was produced by the four participants in which CVC was described in a detailed step-by-step fashion. The document included pictures of all the structured steps of CVC on a real patient, providing a visual example corresponding to the text description of each step. In a final meeting, the final document was approved by all parties. In this same meeting a senior resident (D.T.), who had accepted to teach the CVC course to the interns was briefed and given the final document to study. The document was distributed to the interns of group course one week before the actual course.

Outcomes and evaluation

The outcomes of the study were two: theoretical knowledge about CVC and technical competence in performing CVC on patients. In order to assess knowledge, the interns of both groups completed twice a 15-item multiple-choice-question test. The test (pretest) was first given before the beginning of the study (during the last week of June) and before any of the interns had gained experience with CVC or taken the course. The pretest served to establish the baseline knowledge of the interns of both groups about CVC. Then, the test (posttest) was repeated by the end of the study and immediately after the intern placed a central venous catheter on a patient. Because each multiple-choice question was scored with one point, the highest possible score was 15.

Competence in performing CVC was evaluated by observing the interns performing CVC in critically ill surgical patients. This part of the study was completed within 2 and a half months from the beginning of the study (by mid-August). An independent observer (K.G.T.) evaluated each intern by completing a checklist while the intern was performing the procedure. The evaluator was a qualified surgeon and had participated as an observer throughout the development of the course. At the time of the evaluation he was masked to the group assignment of each intern. During evaluation, a senior resident was present to ensure patient safety but interfered only if the intern had obvious difficulty at any stage of the procedure. The checklist consisted of two fields. One field outlined the 14 important steps necessary to complete the procedure; a “yes” or “no” answer was only possible for any of these items, i.e., the intern did or did not perform correctly each step. Each “yes” answer was scored 1 point, each “no” was scored 0. Therefore, the highest
possible score was 14, which meant a flawless CVC. The second field consisted of 5 additional items evaluating the competence of the intern in CVC with regard to attempts needed to locate the vein and time taken to complete the procedure.

Because the majority of these critically ill patients were receiving sedation and analgesia at the time of CVC, there was no attempt to measure pain and patient discomfort. The safety of CVC was also monitored by recording complications related to it, such as pneumothorax, arterial puncture, significant bleeding, and so forth.

Finally, a supplemental analysis evaluated possible differences between performing CVC on models and patients. For this purpose at the end of the surgical skills laboratory course, after the 12 interns of group course had finished practicing CVC, they were formally evaluated by the standardized checklist while performing CVC on last time on an artificial model. Their performance on the model was compared with their performance on a real patient.

### Statistical methods

The proportions of correct answers in the multiple-choice-question tests and the checklist were compared between the two groups by Fisher’s exact test. For the 15 multiple-choice questions answered in the pretest and the posttest, the P values were adjusted for multiplicity by the stepdown permutation method. The mean values of the scores achieved in each test and the checklist were compared between the two groups by the t test. For the 15 multiple-choice questions of the pretest and posttest, the P values were adjusted for multiplicity by the stepdown bootstrap method. To compare the pattern of answers from the pretest to the posttest, the McNemar’s test was applied. In order to compare the two groups with regard to the second field of the checklist the Wilcoxon two-sample test was used. Additionally, we calculated the ratio of the percentage of “yes” answers on the checklist (ie, steps done correctly) of the course group to the percentage of “yes” answers of the traditional group. This ratio was called “relative likelihood of success” for lack of a better term, and its 95% confidence intervals were calculated. Because the first part of the checklist, describing the steps of the procedure, was newly designed and not used before, a Cronbach’s alpha was calculated for it. Statistical significance was defined by a P value of less than 0.05 for all comparisons. The statistical analysis was performed with the SAS system (release 8.2; SAS Institute, Cary, North Carolina). The study was approved by the Institutional Review Board of our medical center.

### Results

#### Comparison of baseline factors

There was no difference between the two groups in the ratio of preliminary to categorical residents or the departments in which the interns belonged. No difference was noted in previous experience with CVC. The central lines were placed either during the day shift (7:00 AM to 7:00 PM) or the night shift (7:00 PM to 7:00 AM). There was no difference with regard to the shift during which CVC was performed (Table 1).

All patients who had CVC as part of this study were in the surgical intensive care unit. There was no difference between the two groups with regard to certain patient characteristics which might affect CVC. Finally, there was no difference in the anatomical sites (internal jugular or subclavian vein) selected for CVC (Table 1).

#### Comparison of intern knowledge on CVC

Although the mean scores of the multiple-choice pretest were not different between the two groups (7.33 ± 1.07 versus 8 ± 2.15, P = 0.94), group course had a significantly
higher mean score than group traditional in the posttest (11.0 ± 1.86 versus 8.64 ± 1.82, P = 0.03). The difference between the pretest and posttest mean scores was statistically greater in group course compared with group traditional (3.67 ± 2.06 versus 0.64 ± 1.98, P <0.001).

Comparison of intern competence in CVC

The Cronbach’s alpha for the 14-item checklist was 0.76. The percentage of interns who performed correctly each one of the steps of CVC is shown in Table 2. The total mean score on the 14-item checklist was significantly higher in the course group compared with the traditional group (12.6 ± 1.1 versus 7.5 ± 2.2, P <0.001). The relative likelihood of success to perform at least 8 steps correctly during CVC was two times higher among course interns (95% confidence intervals: 1.2, 3.4, P = 0.006). Additionally, the interns of group course required less needle punctures to locate the vein and less direction from the supervising senior resident (Table 2). There was a trend toward less time required to complete the procedure for the course group compared with the traditional group but this difference did not achieve statistical significance (15.4 ± 9.5 versus 20.6 ± 9.1 minutes, P = 0.149). Two complications were recorded, one pneumothorax and one arterial puncture. They both occurred in the traditional group.

Comparison of CVC in models and patients

The performance of the 12 course interns during CVC on an artificial model (as evaluated by the standardized checklist) was compared with their performance during CVC on a real patient (Table 3). There were no significant differences in any of the steps or the total score between the two evaluations.

Comments

Teaching technical skills is a challenging task both for the teacher and the learner. It occurs in a stressful environment, frequently with little preparation, and not unusually at significant risk for the medium of the learning experience, the patient. Despite these downsides the traditional “see one, do one, teach one” method has become the standard of reference for learning how to perform surgical procedures and the main method by which surgeons around the world have been taught how to operate.

The emergence of the concept of training into surgical skills laboratories has revolutionized the methodology of technical skills training. By using animals, inanimate models or computer-generated simulation, trainees can practice
certain tasks in a stress-free, risk-free environment [4,5,17,18]. Because of ethical considerations associated with practicing on animals, technological imperfections related to simulators, and cost issues related to both, training on inanimate models is the preferred method for teaching technical skills in surgical skills laboratories at this point [1].

Even if the concept of surgical skills laboratory-based teaching is appealing, there is little proof that it is better than the traditional method. The effectiveness of surgical skills laboratory courses is based on questionnaires distributed to the trainees, who reported that they were pleased with the course or felt confident that their skills were improved after the course [3,5]. Many related studies are purely descriptive with minimal or no assessment of the effectiveness of surgical skills laboratory courses [3,8,9]. It is obvious that more data must be accumulated to convince medical school administrators that the substantial cost of a surgical skills laboratory is justified.

Providing a quiet and nonintimidating environment to practice technical skills is not enough. A surgical skills laboratory course will potentially be ineffective, if all it does is to allow trainees to perform a technique in artificial models rather than patients in the same nonstandardized and poorly structured fashion, as done during the traditional methods of teaching skills. To achieve optimal effectiveness, a technical skill should be taught in a detailed, step-by-step, standardized, analytical fashion that allows in-depth comprehension of the essential elements of the technique. Additionally, surgical skills laboratory courses should not rely exclusively on long time commitments by attending physicians.

Cognitive task analysis is an educational methodology that satisfies these requirements and stands out probably as an ideal method to teach technical skills [12,13]. It has been used with success by the army, corporations, or health care businesses [12–16]. It also reduces the time needed by an expert to teach the procedure because an intermediate expert can teach it equally effectively. An expert is often unable to articulate why certain steps in a certain procedure are necessary because the procedure has become automated and no longer accessible to conscious reasoning. By cognitive task analysis these automated elements are captured and taught effectively from experts to novices.

Cognitive task analysis involves interviewing at least two advanced experts. The dyadic structured interviews initially have the role of identifying a representative sample of domain problems that need to be solved. A focused and detailed description of each technical step is generated by one expert and then validated by the second expert. The interviews are recorded in a systematic fashion that is self-explanatory and easily understandable. These documents can be studied by the trainees prior to the course to obtain a basic fund of technical knowledge and learn a standardized method to perform the procedure. The advanced experts do not necessarily need to participate in the course and can be substituted by intermediate experts who have participated in the development of the documents and will teach based on these.

In our study, we decided to evaluate the effectiveness of an inanimate surgical skills laboratory course by focusing on the ultimate outcome, i.e. the performance of trainees on humans. We explored the role of cognitive task analysis in teaching a relatively simple technical skill, CVC. After the initial time investment by the advanced experts to develop the course, the training was done by an intermediate expert (senior resident). The course was given on the first week of the interns’ arrival to the Department in order to avoid bias introduced by exposure to CVC during their clinical duties. Evaluation of the trainees was done on knowledge and technical competence. The evaluation was done within a short period of time after the course was taught in order to avoid cross-contamination of the groups to the extent that it was possible.

The results of this study prove the effectiveness of the surgical skills laboratory course. The interns who were randomized to receive the course improved their knowledge about CVC significantly. Although their baseline knowledge was similar, as shown by the scores of the pretest, the course interns scored significantly higher than the traditional interns on the posttest. More importantly, they became more competent technically than the traditional interns, as shown by their scores on the 14-item checklist used to evaluate their performance during CVC. The course interns were more confident than traditional interns, needed less direction by their senior resident, and showed a trend for less time required to complete the procedure successfully. Additionally, they did not cause any morbidity during CVC. Two complications were recorded in the traditional group but conclusions cannot be drawn about safety based on these small numbers. All these findings combinedtranslate to a higher level of competence in CVC by the interns who received the surgical skills laboratory course. An interesting side-finding of this study is that the scores of the course interns on the 14-item checklist were not different when evaluated on a model or on a patient. This shows that they retained their competence gained during the course, even when asked to perform the task in a stressful clinical environment up to two and half months later. This finding is in accordance with existing literature supporting the effective transfer of technical skills from inanimate to human models [19].

To our knowledge this is the first study that provides class I evidence about the effectiveness of a surgical skills laboratory course, as evaluated by a pertinent clinical outcome. Despite its randomized design the study is associated with several limitations. First, it does not have long-term outcome assessments. We did not evaluate many months later if the interns retained the knowledge and competence acquired in the course. It is not likely that differences persist for such a straightforward and frequently used procedure because, as time progresses, traditional intern's are expected
to gain experience through ongoing practice. Second, the instruments of evaluating competence in performing CVC were crude. Psychomotor skills evaluation was not used[20]. Global rating scales—which has been shown to be more accurate than task-specific checklists—were not used either[21]. We believe that a global rating scale is more appropriate for advanced surgical tasks, where the economy and flow of motion can be significantly different from person to person; CVC would not allow for significant differences to show. Third, we had only one evaluator and therefore, assessment of interobserver variability was not possible. The addition of a second evaluator was not possible for logistical reasons that had to do with the “24/7” availability of this person to evaluate an intern. The relatively straightforward steps recorded in the checklist did not allow for a great deal of subjectivity and variability. Fourth, although the evaluator did not know each intern’s group assignment, true masking was impossible. Finally, this study describes a relatively complex design for a relatively simple procedure. Our objective was to develop a valid method to teach technical skills and prove that this method results in measurable differences in the clinical setting. In order to achieve this goal at our first attempt, we needed to avoid multiple variables that could possibly influence outcome. Clearly, an advanced surgical procedure would introduce a high level of complexity and bring along many unforeseeable factors affecting outcome. Now that we feel confident that the method of cognitive task analysis in a surgical skills laboratory setting is valid, we believe that more complex surgical procedures can be taught and result in even greater measurable effects. We are already planning to apply our methodology on advanced surgical procedures.

In summary, we have shown that a surgical skills laboratory course improved significantly the knowledge and skills of the course’s participants in CVC compared with their peers who were trained by traditional methods. This improvement was measured in a real clinical scenario. Cognitive task analysis is a promising educational technique to teach surgical skills laboratory courses because it allows intermediate experts to train novices in a standardized and effective fashion.

Acknowledgments

Supported by a CESERT grant from the Association for Surgical Education.

References