The Use of Cognitive Task Analysis and Simulators for After Action Review of Medical Events in Iraq

Richard E. Clark, Ed.D.*
Carla M. Pugh, M.D. Ph.D.**
Kenneth Yates, Ed.D.*
Maura Sullivan, Ph.D.***

June 21, 2008

Abstract
This study tested a novel approach to medical after action reviews (AAR) by employing cognitive task analysis (CTA) methods. Ten trauma surgeons, who were either deployed in a Forward Surgical Team (FST) in Iraq or at an urban trauma center, were interviewed separately and asked to describe how to perform an emergency shunt procedure to temporarily restore blood flow to the femoral artery. Of the ten surgeons interviewed, nine provided an unaided (no CTA) description of the shunt procedure, and one was interviewed using CTA methods. When compared to a gold standard surgical protocol, the CTA interview resulted in greater accuracy and completeness, whereas descriptions without CTA omitted nearly 70% of the steps. Further, those surgeons who used equipment while describing the procedure gave more accurate and complete reports of the procedure than those who did not. The study provides implications for policies related to after-action review and surgical training.

Introduction
The delivery of health care on the battlefield is managed using a system with escalating levels of capability (USAISR, 2007). The first of five levels is the Battalion Aid Station, which provides triage, treatment, and evacuation. In current conflicts, especially in the early stages of the Iraq conflict, forward surgical teams (FST) co-locate with a Level II capability. One example is a FST with a brigade aid station, which is one terrain feature away from the forward edge of the battle area. FSTs provide surgical resuscitative care for injuries that would not survive a prolonged evacuation process. Typically, an FST consists of three general surgeons, one orthopedic surgeon, and approximately 16 nurses and technicians.

When an FST receives a patient with femoral artery disruption, the medical team is faced with several treatment options (Surgeon 1, personal communication, April 1, 2005). The first

---

1 This Report draws on a work product developed by Dr. Richard Clark of the University of Southern California and submitted to satisfy contract W81XWH-04-C-0093 from U.S Army Medical Research and Materiel Command to the Center for Cognitive Technology, Rossier School of Education. The views, opinions and/or findings contained in this document are those of the author(s) and not those of the University of Southern California, and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
2 *Center for Cognitive Technology, Rossier School of Education, University of Southern California: www.cogtech.usc.edu
3 Additional levels include Level III, a combat support hospital, Level IV, a fixed facility outside of the United States, and Level V is a fixed facility in the CONUS.
option is definitive vascular repair; however, the battlefield conditions and the mission of the FST generally require that surgeons either ligate the vessel and evacuate the patient, or shunt the injured vessel to stop the bleeding and restore blood flow to the extremity prior to evacuating the patient to the next level of care. Whether the patient can be evacuated in a timely manner depends on the evacuation capabilities in the battlefield and the level of activity at the Level III echelon. Thus, the most viable option in an FST is to use a shunt as the highest level of care for these patients.

Army surgeons constantly work to improve their level of technology and training sophistication in the deployed environment. Toward that end, surgeons undergo significant training prior to deployment on how to assess, manage, and treat severely traumatized patients (Defense Medical Readiness Training Institute, 2007). One component of this training includes situational planning and learning how to work with limited equipment. Moreover, severe time constraints as well as battlefield conditions during treatment of patients with serious vascular injuries to the extremities can result in poor workmanship and often results in morbidity or mortality. Situational planning and training for these conditions includes the use of Argyle-type shunts to restore temporary blood flow as a result of damage to the femoral artery.

The Army has adopted the After Action Review (AAR) as the primary method for providing military historical research for training development and performance feedback during training (Morrison & Melliza, 1999). During an interactive discussion, three questions are addressed: (1) “What happened?” (2) “Why did it happen?” and (3) “How can units improve their performance?” As a collective self-examination, however, AARs largely rely on the self-report of participants, and is, therefore, heavily memory-dependent.

In the practice of surgery, an understanding of the conditions and procedures that led to a specific outcome provides opportunities for critical review and improvement over time. Ideally, some record of the surgery is made in real time to allow a complete picture to be developed for review. However, in FSTs such real-time records are either unavailable or impractical.

Studies from the field of cognitive psychology suggest that the use of standard self-report or interview protocols to extract descriptions of events, decision making and problem solving strategies can lead to inaccurate or incomplete reports (Glaser et al., 1985; Besnard, 2000). These errors are not often recognized by those who perform in emergency situations and wish to give accurate reports on their solutions because of the automated and unconscious nature of the knowledge described (Wheatley & Wegner, 2001). This is complicated by the fact that experienced medical personnel mistakenly believe that their reports are complete and accurate and that the problems they are describing were solved in a conscious and deliberate manner (Wegner, 2002). These reporting errors are likely to increase in number and impact under stressful battlefield situations (Hunt & Joslyn, 2000).

Cognitive Task Analysis

CTA uses a variety of interview and observation methods to capture the knowledge, goals, strategies, and decisions that underlie observable task performance (Clark, Feldon, van Merriënboer, Yates, & Early, 2007). CTA can take many forms, but in all cases seeks to capture the knowledge of subject matter experts (SMEs) who have demonstrated consistent proficiency in performing a task over the long period of time. The five common steps found in most of the dominant CTA methods are performed in the following sequence:

1. Collect preliminary knowledge to become familiar with the domain and to identify the tasks and procedures being analyzed.
2. Identify knowledge representations of each sub-task and the types of knowledge required to perform it.
3. Apply focused knowledge elicitation methods with subject matter experts to elicit the knowledge identified to perform the task.
4. Analyze and verify data acquired by reviewing transcripts and using multiple SMEs.
5. Format the results for the intended application, such as rules of thumb, gold-standard protocols, and job aides.

Cognitive task analysis has been applied to a wide range of skills, from fighting fires (Klein, 1989) to performing surgical procedures (Dominguez, Hutton, Flach, & McKellar, 1995). The common goal of CTA is to assist a subject matter expert in the retrieval and recounting of a procedure that may be highly automated, and therefore not generally available for conscious inspection.

Most forms of cognitive task analysis show tangible benefits with respect to the accuracy and completeness of knowledge obtained. For example, the use of CTA methods with one expert has been demonstrated to provide a 28% increase [from 12% (Chao & Salvendy, 1994) to 40% (Clark & Estes, 1996)] in the amount of information captured from experts during performance of a task. When multiple experts are interviewed, the percent of information captured increases proportionately (Chao & Salvendy, 1994).

Studies also provide evidence for the efficacy of CTA-based surgical instruction. For example, in a study with first-year medical interns, an expert surgeon taught a procedure for inserting a central venous catheter in a lecture/demonstration/practice sequence (Maupin, 2003; Velmahos et al., 2004). For the treatment group, the lecture was generated using CTA methods with two experts in the procedure. The control group’s lecture consisted of a traditional instructional practice in which the expert instructor engaged in a free recall explanation of the procedure while providing a demonstration. Students in both conditions were provided equal time for questions, practice, and access to equipment, and they also completed a written posttest and performed the procedure on multiple human patients during their internships. The results showed that students in the CTA condition had significantly greater gains from pretest to posttest than those in the control condition. Moreover, they outperformed the control group in actual patient practice in every measure of performance, including an observational checklist of steps in the procedure, number of needle insertion attempts needed to insert the catheter into patients veins, frequency of required assistance from the attending physician, and time-to-completion for the procedure.

**Hypotheses:**

The purpose of our study was to determine whether a combination of CTA methods and simulations could improve AAR by more accurately capturing the automated and conscious decisions surgeons make as they describe the procedure for inserting a femoral artery shunt to restore blood flow to a damaged extremity under extreme conditions. We hypothesized that:

1. Surgeons who give unaided description of the surgical protocol will omit ± 70% of the critical steps in the surgical procedure, when compared with a gold standard protocol;
2. The completeness of unaided, self-reported surgical protocol information will vary within different segments of the surgical protocol;
3. The increased use of technology during description will increase the accuracy and completeness of surgical protocols; and
A CTA interview will increase the accuracy and completeness of surgical protocols by between twelve and forty percent.

This application of CTA involved asking a series of questions designed to cover all aspects of the medical procedure – from the conditions under which the surgery is indicated, to the materials and/or equipment needed. In addition, we inquired about action and decision steps, performance standards, concepts, processes, principles underlying the procedure and finally completion. Lastly, we thought that the availability of surgical equipment and visual aides during CTA might increase the accuracy of the protocols by stimulating the recall of critical surgical steps (Pugh & Clark, 2006).

Method

Task: The task targeted for this study – insertion of a femoral artery shunt – was selected because it is an increasingly common procedure for surgeons working in Iraq, and because it has not previously been captured using CTA. When a forward surgical team (FST) in a battlefield context receives a patient with femoral artery disruption, the medical team is faced with several treatment options. The first option is definitive vascular repair; however, the battlefield conditions and the mission of the FST generally require that surgeons either ligate the vessel and evacuate the patient to the next level of care or shunt the vessel to restore adequate blood flow to the extremity. Whether the patient can be evacuated in a timely manner depends on the evacuation capabilities in the battlefield as well as other factors such as availability of treatment elsewhere. Often, the most viable option in an FST is to use a shunt to restore blood flow pending definitive vascular and/or orthopedic repair at a more completely equipped care facility.

Sample: A total of eleven trauma surgeons were interviewed for this study. Nine trauma surgeons with experience in placing vascular shunts in emergency trauma environments were recruited from the medical school of a large urban research university. A tenth trauma surgeon who had experience performing shunt placements as part for a FST in Iraq was recruited through the offices of the study sponsor. The eleventh trauma surgeon, also recruited through the university, had experience in both urban and battlefield settings, was recruited to complete and approve the gold standard protocol.

Design

Nine trauma surgeons, who have used Argyle-type shunts to repair femoral artery damage in an urban environment, were interviewed. The surgeons provided an unaided description of the procedure (“no CTA” condition). Follow-up questions were asked to clarify statements made in the unaided section of the interview. Each surgeon was interviewed separately and asked to describe how to perform the shunt procedure under emergency conditions. Five surgeons described their surgical protocol as they manipulated a set of surgical instruments and viewed depictions of the anatomy surrounding the femoral artery. The remaining surgeons were interviewed without access to instruments or anatomy and comprised the “no technology” condition.

A full CTA was conducted with a tenth trauma surgeon, who had recently returned from deployment in a Forward Surgical Team in Iraq. A trained interviewer, who was not a surgeon, conducted the semi-structured interview using the protocol attached as Appendix A, which consists of a series of questions pertaining to (1) Conditions (indications and contraindications for performing the procedures) (2) Processes (who does what, when, and where); (3) Steps

These questions as well as an explication of the CTA model implemented in this study are found in Appendix A.
(action steps and decision steps accompanied by alternatives and the criteria for deciding); (4) Standards (time and quality); (5) Equipment (instruments); (6) Reasons (principles of science, i.e., why do this, and not that); and (7) Concepts (names, symbols, or events that a surgeon would need to know to perform the procedure).

All interviews were recorded with audio and video. The interviews were transcribed verbatim, and were coded using a scheme based on the interview questions. Two coders worked independently to code the transcripts, and then met to compare the results and resolve the discrepancies. An inter-rater reliability of .87 was achieved in the coding.

A protocol of the each surgeon’s description of the procedure was developed from the coded transcripts by one coder and reviewed for accuracy by the other coder. The surgeons were then given the opportunity to review and correct the protocol developed from their transcribed interview. The surgeons’ corrected protocols were aggregated to create a preliminary “gold standard” protocol. An independent vascular surgeon, who was a senior member of the faculty at a leading urban medical school, reviewed and corrected the preliminary protocol, which then became the final “gold standard.”

As the final step in data gathering, the accuracy and completeness of each surgeon’s interview data, represented by the statements contained in the protocol, were compared to the gold standard to determine the gain or loss of AAR fidelity due to CTA use with and without simulators. Surgeons’ protocols were compared and analyzed prior to being corrected (Round 1), and after review and correction (Round 2). Additional data included the level of experience, measured by the surgeons’ report of the number of shunt procedures performed, and a review of the video tape record to assign the level of technology interaction during the interview, for those surgeons in the technology condition. A rating scale of technology (use of instruments and pictures of anatomy) was developed so that: “Minimal” referred to a surgeon’s reference to technology verbally, visually, or by pointing; “Occasional” referred to a surgeon’s occasional touching of the technology; and “Heavy” referred to a surgeon’s use of the technology to demonstrate the procedure.

Results

The study was designed to test four hypotheses.

Hypothesis 1. Surgeons who give unaided description of the surgical protocol will omit ± 70% of the critical steps in the surgical procedure, when compared with the gold standard protocol. The total percentage of agreement between surgeons’ description of the shunt procedural steps and the gold standard protocol in the unaided interview condition were 25.00% in round one with a 6.25% improvement in round two for a total of 31.25% agreement. Thus, surgeons omitted 68.75 % of the standard procedural steps in support of Hypothesis 1.

[HYPOTHESIS 1 GRAPH HERE]

Hypothesis 2. The completeness of unaided, self-reported surgical protocol information will vary within different segments of the surgical protocol. As shown by Figure 2, the percentage of agreement with the gold standard protocol varied within the surgical protocol. Each of the codes varied from the average level of agreement which supports hypothesis 2.
Hypothesis 3. The increased use of technology during description will increase the accuracy and completeness of surgical protocols. Figure 3 indicates that there is no difference between the Minimal and Occasional use of technology in the accuracy and completeness of recall when compared with the gold standard. However, heavy use of technology increased the level of agreement in both Round 1 and Round 2. These data may support the hypothesis that the increased use of technology augments surgeons’ description of the procedural steps. It also appeared that those surgeons who made use of instruments and pictures during the interview experienced greater recall.

Hypothesis 4. A CTA interview will increase the accuracy and completeness of surgical protocols by between twelve and forty percent. As shown in Figure 4, a comparison between the no CTA and CTA conditions demonstrates the increase in the percentage of agreement between aided and unaided descriptions of the shunt procedure when compared with the gold standard. The results show that the total percentage of agreement with CTA is 68.75% compared with 31.25% without CTA in support of Hypothesis 4.

Conclusions and Implications

This study demonstrates that consideration should be given to adopting CTA for critical AARs to increase accuracy and decrease recall errors. Training curricula based on unaided (non CTA) expert interviews thus, incomplete information on task performance, may lead to less effective training. Moreover, simulation technology developed using unaided interviews may also be less effective in meeting the learner’s needs. To address the above issues, it may be desirable to use “gold standard” protocols developed through the CTA process for surgical skills training and assessment. More studies should be conducted that compare CTA-based training curricula with traditional Halstedian “see one, do one, teach one” pedagogy (Halsted, 1904). This project sought to replicate the AAR environment to demonstrate that when CTA methods are applied, both the accuracy and completeness of surgeons’ descriptions improve. These results hold promise for the use of CTA in combination with AARs to enhance “lessons learned” and training effectiveness and efficiency, when contrasted with the use of AARs alone.

Future Directions

The application of CTA methods with subject matter experts (SME), produce high quality representations of expert performance in increasingly complex and difficult settings. This information can be developed into simulations that provide efficient and effective practice of surgical skills acquired during training. Simulators provide an environment for surgeons to repeatedly and deliberately practice new surgical skills while receiving immediate and corrective feedback, based on desired expert performance captured by the CTA enterprise. Current surgical simulators largely focus on technical skills. Future research should focus on the use of CTA to...
expose critical decision-making skills necessary to perform difficult operations and surgical tasks. Incorporating “gold standard” CTAs as the basis for assessing technical and surgical decision making skills should also be examined.

It is also important to realize that all CTA interview protocols are not equally effective (Clark, Feldon, vanMerrienboer, Yates & Early, 2007). Based on the ultimate use of the CTA outcomes, knowledge representations are identified early in the CTA process. These representations determine the various CTA techniques that should be used to elicit the appropriate type of knowledge. Some methods are better suited to capture procedural knowledge while others are more effective to elicit declarative knowledge (Chipman, Schraagen, & Shalin, 2000). Expert systems or computer-assisted tutoring applications, for example, require CTA methods that capture highly structured procedural knowledge. Whereas, knowledge representations that support training applications require both declarative and procedural knowledge. Less formal CTA methods, such as the one described here, are generally more effective for capturing the both the underlying conceptual knowledge and procedural skills experts use while performing complex tasks.

References


Figures

Figure 1: Total Percentage of Agreement Between Surgeons Unaided Description of Shunt Procedure and Gold Standard Protocol
Figure 2: Percentage of Agreement of Unaided Surgeons Description of Procedural Steps by the Level of Interaction with Technology
Figure 3: Total Percentage of Agreement Between Surgeons Description of Shunt Procedure with CTA and without CTA - when compared to Gold Standard procedure.
Figure 4: Total Percentage of Agreement Between Surgeons Aided and Unaided Description of Shunt Procedure and Gold Standard
APPENDIX A: Surgical CTA Protocol

1. Establish general context for use of procedure, general indications and contraindications for use and any relevant history (see “Questions to be asked…” below, items 1-3).
2. Ask for a sequential explanation of the process. Emphasize that instructions should be given as they would to an intermediate medical student, and request that steps be described as specifically and completely as possible, including decisions that must be made, cues that must be attended to, etc. Remind the subject to use ordinal descriptors as frequently as possible (e.g. “First, do step 1. Second do step 2. Next, do step 3…”). Questions/clarifications should only be asked of the subject if the words used or pronoun-antecedent relationships are not clear. Ask also about the decisions that must be made and the criteria for choosing between the various alternatives when decisions are made (see “Questions to be asked…” below, items 4 and 5)
3. Recite the sequence back to the subject – that is, paraphrase what you hear them say. Ask for corrections and clarifications.
4. Ask the subject if the sequence after the corrections and clarifications is sufficient to allow someone to complete the task successfully (see “Questions to be asked…” below, items 6-7).
5. Take a break. Compile notes into a single, step by step, action and decision procedure.
6. Ask the subject to listen to you talk through the procedure as if someone was performing the procedure in a hypothetical situation. Instruct him to interrupt, clarify, or correct if anything said is inconsistent with how he/she would perform the procedure.
7. Review the corrected procedure with subject. At each identified decision point, ask for all relevant cues (see “Questions to be asked…” below, items 4-5). Verify by rephrasing as a question (e.g. “So, in order to make this decision, I only need to look at these two things?”).
8. In preparation for a follow-up meeting, compile the written CTA document and send to the subject. Ask him/her to make any changes that are necessary to correct the accuracy of the CTA using the “track changes” function in Microsoft Word or to print a copy and bring handwritten notes for changes to the follow-up meeting.
9. At follow-up meeting, discuss all changes and finalize the CTA description. Explain to subject that when he is asked to review others’ CTA documents, his role is to determine whether or not the task can be successfully completed using the steps presented. He should neither assume that something unstated is known nor that the CTA should exactly match his personal procedure. He should also edit any steps that are unnecessary. The emphasis should be on whether or not the CTA document to be reviewed is viable and efficient to complete the task as written. Any changes that the subject wants to make should be made in the same manner as the edits to his own document.

Questions to be asked during the interview protocol

1) What happened? What were the problems being solved and the medical goal of this event? The objective of this question is to collect the expert’s overview description of the “what, where, when, who, why” the event happened. In addition, background information on the precursors, context, preparedness, important and unexpected aspects of the event are collected as well as the expert’s view of the goal to be achieved.
2) *What conditions must be present to start the task?* Here the goal is to collect information about the medical “conditions” or “indications and counter indications” that would permit medical personnel who have not experienced this event to know when it has occurred and how to identify it unambiguously. Any tests, observations or measurements that must be made are collected and described.

3) *What is the reason for the unique or unexpected nature of this event?* The goal here is to collect background information on why this event was perceived as unexpected or important. The interviewer usually asks what aspect of prior training or education prepared the expert for this event – and what might prepare future surgeons more adequately to deal with it.

4) *What actions and decisions must be implemented to complete the task? What alternatives must be considered and what criteria must be used to decide among the alternatives?* This question is the core of a CTA interview. The expert is asked to describe, in a step-by-step fashion, everything that must be done to diagnose and treat the problem being investigated. This is often the second question that is asked (after #1, “what happened”). The answers to questions # 2 and 3 most often turn up as the expert describes the sequence they followed to diagnose and treat. As the sequence unfolds, the interviewer often interrupts with questions about the actions being described such as “Can you demonstrate on the simulator what you are describing?”, or “Why did you do that?”, or “What alternatives did you consider and what criteria did you use to make that decision?” and “What would lead you to make a different decision with another patient? Could you demonstrate a different set of constraints for that decision on the simulator?” The key issue in a CTA is to capture all of the many complex decisions that must be made, the alternatives that must be considered before a decision is reached and the essential criteria for choosing between the alternatives. It is knowing when and how to make decisions that are most often the source of errors in medical training since experts tend to automate their decision making. While experienced experts make very rapid and accurate decisions, they cannot observe what goes on in their mind as they decide and so often fail to report decisions or the range of alternatives they considered and rejected. This information contributes to training that is often very accurate when it depicts the observable actions that subject matter experts (SME) use to solve problems but unobservable decisions are often ignored or distorted. The goal of this aspect of the CTA is to produce an accurate, step by step description of the most efficient and effective way to reach the medical goal and sub-goals of the task.

5) *What concepts, processes or principle knowledge is required to adjust this task to fit novel conditions?* As the expert describes actions and decisions in response to question #4, the CTA interviewer occasionally interrupts and asks for details about three types of knowledge. A) Concepts -- An explanation of the special medical or scientific terms used by the expert. The interviewer asks for definitions and identifiable examples. Examples are collected (and scanned or otherwise stored on a computer for later use as illustrations in the CTA). Concepts are the type of knowledge that supports accurate classification of all aspects of the problem and solution. B) Processes -- An explanation of how something important to the goal works, stage by stage -- such as a disease progression or an organ system. Processes support clear understanding of the wider context of the systems involved in the problem and solution and help experts generate more adequate solutions to problems; and C) Principles - Essentially the “science” of the
phenomenon being described in the form of variable cause and effect statements. Principles help identify and explain causes, solutions and the adjustment of procedures to accommodate highly important incidents related to the problem being studied. These three types of knowledge will eventually be reorganized and presented as the body of conceptual and scientific knowledge that will support the diagnosing and treating of the problem and the editing of established treatments to accommodate unusual cases.

6) What equipment and materials are required? The objective with this question is to determine if any unusual medical equipment or supplies, not usually available in the context where this problem might occur, need to be provided in order to effectively diagnose and treat the problem effectively. Descriptions of equipment are collected and scanned or stored on a computer for later use in the CTA report.

7) What performance standards must be achieved? (E.g. time, accuracy). All essential quantity and quality standards for the diagnosis and treatment of the problem must be identified so that they can be described in assessment instruments and for eventual training media and materials.