Mathan, S. & Dorneich, M. (2005). "Augmented Tutoring: Improving military training through model tracing and real-time neurophysiological sensing", Foundations of Augmented Cognition (D.D. Schmorrow, Ed.). Mahwah, NJ: Lawrence Erlbaum Associates. pp. 964-973.

Augmented Tutoring: Enhancing Simulation Based Training through Model Tracing and Real-Time Neurophysiological Sensing

Santosh Mathan, Michael Dorneich

Honeywell Laboratories
3660 Technology Dr, Minneapolis, MN 55418
santosh.mathan@honeywell.com

Abstract

Military training simulations provide rich and engaging environments for personnel to develop and maintain mission critical knowledge and skills. However, these systems generally lack the capacity to diagnose student difficulties in real-time and provide automatic, context specific assistance to learners. The lack of pedagogical diagnosis and guidance has several implications for the effectiveness and efficiency of the training process. Without feedback, students can take problem solving paths that deviate far from the solution. Besides introducing inefficiencies, the lack of appropriate guidance can contribute to unproductive floundering and induce confusion and frustration among students. In this paper we describe a set of diagnostic technologies that could provide the basis for inferring both a student's immediate problem solving context and underlying cognitive state. Dynamic guidance within simulation environments based on such a comprehensive assessment of student state could serve to raise both the efficiency and effectiveness of military training simulations.

1 Introduction

Computer based military simulations have emerged as cost effective and engaging mediums for military personnel to develop and maintain operational proficiency. They range in sophistication from embedded training systems incorporated within deployed equipment (e.g. Aegis Combat Training System), to computer games designed with guidance from military domain experts. These systems allow students to practice tasks in environments that bear a high degree of fidelity with real world task contexts. Unfortunately, a key element necessary for achieving the promise of effective ubiquitous training is largely missing in most military training simulations. These systems typically lack the ability to scrutinize student performance and provide context specific guidance to learners in real-time. This could contribute to sub-optimal learning outcomes.

One strategy for assessing student expertise within training simulations is to examine success at accomplishing mission objectives. Unfortunately, global performance outcomes only provide a coarse indication of a student's competence. In order to maximize training effectiveness, it is critical that the training environment assess performance at a finer grain size. In complex task contexts, a broad range of strategies and tactics combine to contribute to the overall performance outcome. For example, consider the case of a unit leader practicing scenarios within a tactical decision simulator. The objective of a particular scenario might be to lead an assault on a building in an urban environment. Besides assessing the ultimate outcome, it may be critical to determine whether the unit leader appropriately coordinated fires — whether a long covered approach selected over a short open route — whether obscuration with smoke was invoked by the student when relevant — whether bounding overwatch maneuvers used when appropriate. Poor performance on one or more of these sub tasks could compromise battle field effectiveness. However, without active monitoring of performance of tactics and strategies that contribute to mission outcomes, many performance lapses may go undetected — potentially denying the trainee an opportunity to acquire and reinforce critical decision making knowledge and skills.

In order to overcome the shortcomings associated with the lack of fine-grained diagnosis and feedback in training simulations, the military has relied on human observers. But such an approach has limitations too. These are highlighted in a study conducted in the context of the Aegis Combat Training System, a ship-based embedded training system (Zachary, Cannon-Bowers, Bilazarian, Krecker, Lardieri, & Burns, 1999). While the ACTS system presents impressive high fidelity simulation capabilities, the lack of functionality to examine student actions in real-time and provide timely remedial feedback, presents numerous difficulties. As Zachary and colleagues have noted, every student in the ACTS environment requires an experienced crew member watching over his or her shoulder to identify difficulties and provide feedback. Unfortunately, these observers are often not training experts. Researchers noted inconsistencies in the quality of help they were able to provide students. Additionally, it was difficult for human observers to analyze and provide feedback on the hundreds of actions that a student might perform in the highly dynamic, simulation environment. Furthermore, the 1:1 student to instructor ratio that Zachary and colleagues noted may be unfeasible in many operational training contexts.

While automated pedagogical diagnosis can raise the efficiency and effectiveness of training simulations, a broad range of technical challenges limit their widespread use. Some of the technical challenges include:

- *Inferring Problem State*: Students interact with dynamic simulation elements in a relatively unconstrained fashion. Learners are capable of employing a broad range of strategies in solving complex tasks. As a consequence, identifying a learner's immediate problem solving state with respect to the problem solving goal is often difficult. The system has to recognize a range of strategies and assess them in terms of their overall effectiveness and efficiency. Additionally, the system must incorporate sufficient knowledge of the domain to provide appropriate context specific feedback.

- *Assessing Evolving Student Knowledge:* In order to optimize the pace at which students progress through a training sequence, it is important that the system be able to assess a student's competence on a potentially large set of knowledge and skill components that may be leveraged to solve problems. Accurate characterization of evolving student knowledge and skill would allow the system to select problems that target specific weaknesses, instead of stepping all students through a canned sequence of problems. Keeping track of a students evolving competence on the broad range of knowledge and skill components necessary to solve problems is a difficult technical challenge

- *Assessing Underlying Cognitive State:* Research suggests that aspects of a student's underlying cognitive state, such as working memory capacity, and attention, have a direct impact on the ability of students to learn. However, assessing these states and tailoring feedback based on these assessments is a difficult challenge and practically never done in the context of computer based learning systems.

This paper presents two powerful technologies — ACT-R based cognitive tutors and non-invasive neurophysiological sensing — that could provide the basis for addressing the technical challenges outlined above. A simulation environment embodying these features would dynamically guide students toward expertise — with feedback based on both an assessment of overt problem solving actions and parameters such as working memory load, attention, and cognitive arousal that have an impact on learning outcomes.

In the following pages we elaborate on each of the technical challenges mentioned above and discuss technical solutions to address each of these problems.

2 Technical Challenges and Promising Solutions

2.1 Tracking Problem State

In complex, scenario-based simulation environments, students have access to a broad range of problem solving actions (operators) that can be combined to transform a problem from some initial state to a goal state through set of intermediate problem states. In complex task domains the number of actions students have at their disposal can be large. By interacting with elements of the simulation environment, these actions can produce a vast set of possible intermediate problem solving states. Certain sequences of problem states may lead to dead ends, other transition paths to the goal state might be inefficient, whereas one or more possible sequence of state might allow learners to get to the goal state efficiently. Navigating unfamiliar problem spaces can be difficult for learners. Dead ends and inefficient paths to the goal can induce frustration and confusion among students; they can also lead to the acquisition of sub-optimal performance strategies.

2.1.1 Model Tracing

A promising approach for boosting the diagnostic capabilities of military training simulations may be found in ACT-R based cognitive tutor technology (Anderson, Corbett, Koedinger, and Pelletier, 1995). ACT-R is a theory of cognition that describes how humans perceive, think and act (Anderson and Lebiere, 1998). ACT-R is instantiated in the form of a programming language whose primitive constructs embody assumptions about human cognition. It has been used to model human performance in a broad array of complex cognitive tasks, ranging from automobile driving (Salvucci, 2001) and tactical decision making (Anderson, Bothell, Byrne, Douglas, Lebiere & Qin, in press) to algebra and computer programming (Anderson et al., 1995). ACT-R's development has been led by psychologist John Anderson and colleagues at Carnegie Mellon University over the course of three decades of research in cognitive psychology and artificial intelligence.

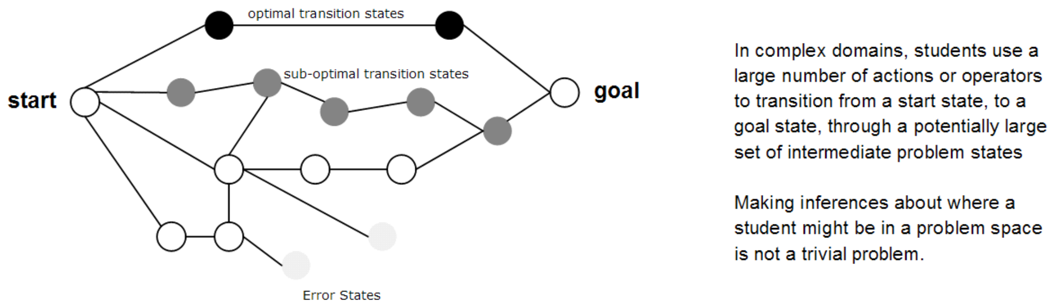

Figure 1: Challenge of tracking students in complex problem spaces

ACT-R-based cognitive models have had particularly broad impact in the area of computer based tutoring. Through a process called model tracing, cognitive tutors assess student actions as they perform complex cognitive tasks using computers. Model tracing is an instance of a plan recognition algorithm — a class of artificial intelligence programs that deal with the issue of inferring an agent's plans or line of reasoning from its actions (Kautz and Allen, 1986).

The detailed encoding of knowledge embodied in a cognitive model allows the model tracing algorithm to pin-point a student's progress through a problem space. When student actions are consistent with one or more fruitful problem solving strategies, the system remains unobtrusively in the background. However, when student performance is consistent with ineffective or inefficient strategies, the system intervenes with assistance that is tailored to the specific difficulty being faced by a student at a given moment. A student may also ask a tutor for hints that are tailored to a learner's immediate problem solving context.

Model tracing can also serve to simplify after action reviews. Generally, logs generated by simulation environments are difficult for instructors to base instructional feedback on. They commonly express a user's actions in terms of screen coordinates, keystrokes, mouse clicks, and hits taken, to name a few parameters. However, information that is useful from a pedagogical point of view is a trace of a student's thought process in relation to events within simulation scenarios. Since the model tracing process maps a student's actions to cognitive steps expressed in the ACT-R cognitive model, it is possible for the instructor to make inferences about student performance on the basis of a student's thought processes rather than depending on low level actions and events in a simulation alone.

2.2 Assessing Evolving Student Knowledge

Problem solving in complex domains relies on a wide range of knowledge and skill components. Global performance metrics gathered at the end of a task shed little light on a student's evolving competence on the knowledge and skill components that must be leveraged for effective task execution. For example, consider the example of a student trying to master the skill associated with putting a golf ball. The successful execution of task requires competence in areas ranging from the ability to judge the slope of the green, to posture and stroke mechanics. If a student is consistently performing poorly, it is important to be able to identify the specific deficiency that may be contributing to the outcome. This would allow the training process to target particular areas of weakness.

2.2.1 Knowledge Tracing

Many cognitive tutors incorporate functionality known as *knowledge tracing* to optimally pace students through problems and distinguish between slips and errors. Knowledge Tracing relies on Bayesian estimation procedures to estimate a student's strengths and weaknesses relative to the knowledge components in the cognitive model (Corbett and Bhatnagar, 1997). These estimates are dynamically updated as a student is performing tasks. Knowledge tracing estimates are used to pick problems in areas that a student may need most practice on. Additionally, students get estimates of their mastery of various knowledge and skill components via an on-screen bar graph. Unlike many computer based environments, which guide all students through a set sequence of problems, knowledge tracing allows students to work on problems that are appropriate to their competence level. Proficient students can progress quickly to challenging problems, while students who need additional practice get to work on problems that target their particular deficiencies.

2.3 Real world efficacy

ACT-R based cognitive tutors have been rigorously assessed in classroom and laboratory contexts. These systems have been shown to reduce training time by half and increase learning outcomes by a standard deviation or more (Anderson et al., 1993). They have been used to teach concepts ranging from programming to genetics, and represent some of the most broadly used educational systems. ACT-R based tutors for algebra and geometry are in use by 200000 students in over 1800 schools around the country. The US Department of Education has designated cognitive tutors one of 5 exemplary curricula in K-12 mathematics education.

2.4 Assessing Underlying Cognitive State

While replications in numerous domains have shown cognitive tutors to be among the most effective computer based learning platforms, their performance falls short of one-on-one tutoring from highly-skilled human tutors. Researchers have argued that the advantage skilled human tutors have over cognitive tutors may stem from the fine-grain access they have to their student's behaviors. For instance, as Anderson and Gluck (1999) have noted, a human tutor can see frustration on a student's face, hear uncertainty in an utterance, and keep track of how long a student is taking to respond to a problem. Such broad access to a learner's emotive and cognitive state allows the skilled human tutor to display far greater sensitivity and adaptability in the tutorial interaction than a computer based tutor.

While cognitive tutors have conventionally relied on an interpretation of overt behavioral actions to make inferences about students' problem solving progress, it is now feasible to incorporate interpretations of data ranging from eye movements and electroencephalogram (EEG) data to electrocardiogram (EKG) readings to make inferences about a student's cognitive state. Researchers have referred to this expanded diagnostic capacity as *high density sensing* of student state. Computational capabilities on ordinary desktop computers make it possible for the grain-size of tutorial analysis to shift from the 1 to 10 second level of analysis of problem solving actions to the millisecond level of analysis of physiological and neurophysiological states. However, the means by which such a capability can be harnessed to boost tutorial outcomes remains a largely unexplored research area.

The DARPA funded Augmented Cognition program has played a pivotal role in furthering the development of non-invasive physiological and neurophysiological sensing to identify states that could negatively impact human performance. Applications of technology developed under the Augmented Cognition program have primarily focused on ways to aid human task performance. Neurophysiological and physiological sensors are used to detect cognitive bottlenecks and invoke assistance aimed at helping users perform tasks effectively under extremely demanding conditions. Assistance has included strategies such as task offloading, task sharing, modality switching, and task scheduling.

Applications of augmented cognition in the context of the training simulations require a slightly different perspective. The primary objective of using augmented cognition techniques in the context of tutoring applications is not to simplify task performance with adaptive assistance. Rather, neurophysiological sensing could be used optimize instructional efficacy by dynamically tailoring the instructional environment to match the cognitive capacities of a student. AugCog technologies can play a useful role in two distinct phases of the instructional process as detailed below.

2.4.1 Augmented Tutoring

Declarative Instruction

The process of acquiring a novel skill typically begins with a period of declarative instruction when a student learns about the central facts or concepts associated with a domain. In computer based learning environments declarative instruction is typically facilitated through video clips or online textual expositions. Research has shown that unless attention is appropriately directed towards the processing of declarative information, the robustness and accuracy of these facts in memory is likely to be compromised (Anderson, 1993). A poor declarative encoding can impede skill acquisition. Research shows that declarative knowledge of the central concepts in a domain serve to structure early problem solving attempts (Anderson, 1993).

Acquiring a robust and accurate encoding of declarative knowledge requires both sustained attention and active elaboration strategies such as self-explanation. Unfortunately, research indicates that learners have a hard time maintaining attention over long periods of time. For instance, Schooler (in press) observed subjects over the course of a 45 minute reading task. Subjects were interrupted at random and asked if they were still on task. His research revealed that learner's "zoned-out" for close to 20% of the time over reading tasks. Such lapses in attention over the course of reading text and watching video expositions can have a negative impact on the skill acquisition process. Learners may miss critical information that could be of importance in subsequent problem solving efforts.

Research suggests that it may be possible to detect and mitigate low attentional states. Prior efforts aimed at detecting low attention levels using neurophysiological sensors have been fruitful. For instance, Jung, Makeig, Stensmo, and Sejnowski (1997) have shown that EEG spectral power in the 4Hz and 14Hz bands can predict low alertness levels. Additionally, cardiovascular measures such as inter-beat intervals, can serve as an indicator of cognitive arousal. Sympathetic and parasympathetic components of the autonomic nervous system that govern cognitive arousal can be tracked using the Clemson Arousal Meter, a classifier that is an important component of Honeywell's current Augmented Cognition efforts.

Mitigation strategies, triggered by cognitive state classifiers could minimize the negative impact of low attentional states during declarative instruction. For instance, if the low attentional states are detected while a student is reading text online or watching video segments, the system could intervene and step the student through the material with interactive prompts. These prompts could present questions related to concepts just covered and give students the chance to respond using multiple choice responses. Additionally, the system could index text or video segments that may have been encountered by the students during low attentional states. The system could prompt students to revisit these segments at a later time.

Diagnosis during hands-on Practice

Diagnosis based on non invasive neurophysiological and physiological sensing could also play an important role during hands-on practice with simulation platforms. Learning is a working memory intensive process. As Mayer (2001) has suggested, learning involves building connections between incoming materials and existing knowledge. This integration occurs in working memory. Several cognitive theories posit that excessive working memory loads during the learning process can interfere with the acquisition of problem solving schemas (e.g. Sweller, 1998). While the negative consequences of excessive working memory load on learning outcomes are well known, few practical techniques exist to tailor the problem solving process based on a dynamic assessment of a learners working memory capacity.

EEG based neurophysiological sensing could provide a way to assess working memory loads in real-time. Researchers have noted that levels of frontal theta activity increase as a function of working memory load; whereas levels of parietal alpha show a corresponding decrease (Gevins, Smith, McEvoy, Yu, 1997). EEG based classifiers can detect these states and drive mitigations to modulate working memory load within learning environments. Additionally gauges used by the Honeywell AugCog Team to detect high cognitive load (including the EEG based Executive Load Index, and the composite EEG-EKG-pupilometry-based Stress Gauge) could be adapted for use within tutoring environments.

Neurophysiological assessments of working memory may be used to dynamically match working memory demands of a learning environment with a student's working memory capacity. For instance, the level of assistance or scaffolding provided to students following errors during practice could be based on neurophysiological assessments of cognitive state. The grain size of instruction could be dynamically varied based on these assessments (c.f. Anderson et. al., 1995). Students experiencing high levels of cognitive load could be interactively stepped through the series of sub-goals necessary to accomplish a problem solving goal. In contrast, students experiencing lower cognitive load levels could simply be reminded of the overall problem solving goal — maintaining and negotiating the underlying goal stacks in memory could be left to the student. Another strategy that could be used to modulate workload levels might be to adapt the pace of the simulation to the working memory capacity of a student. As a student becomes more proficient at performing tasks, the working memory resources associated with task execution can be expected to diminish.

3 System Implementation

Several elements are necessary for the system to provide feedback based on the joint consideration of cognitive state and problem solving context. First, a simulation environment has to be instrumented to communicate information about simulation objects, events and student actions to the underlying cognitive tutor. Second, an ACT-R cognitive model characterizing the cognitive steps associated with strategies learners are likely to adopt has to be developed for the instructional tasks. Third Information from neurophysiological sensors has to be filtered and classified in terms of cognitive states of interest. Feedback and other mitigations have to be tailored to leverage cognitive state and problem state assessments. Figure 2 shows a possible architecture for an augmented tutor.

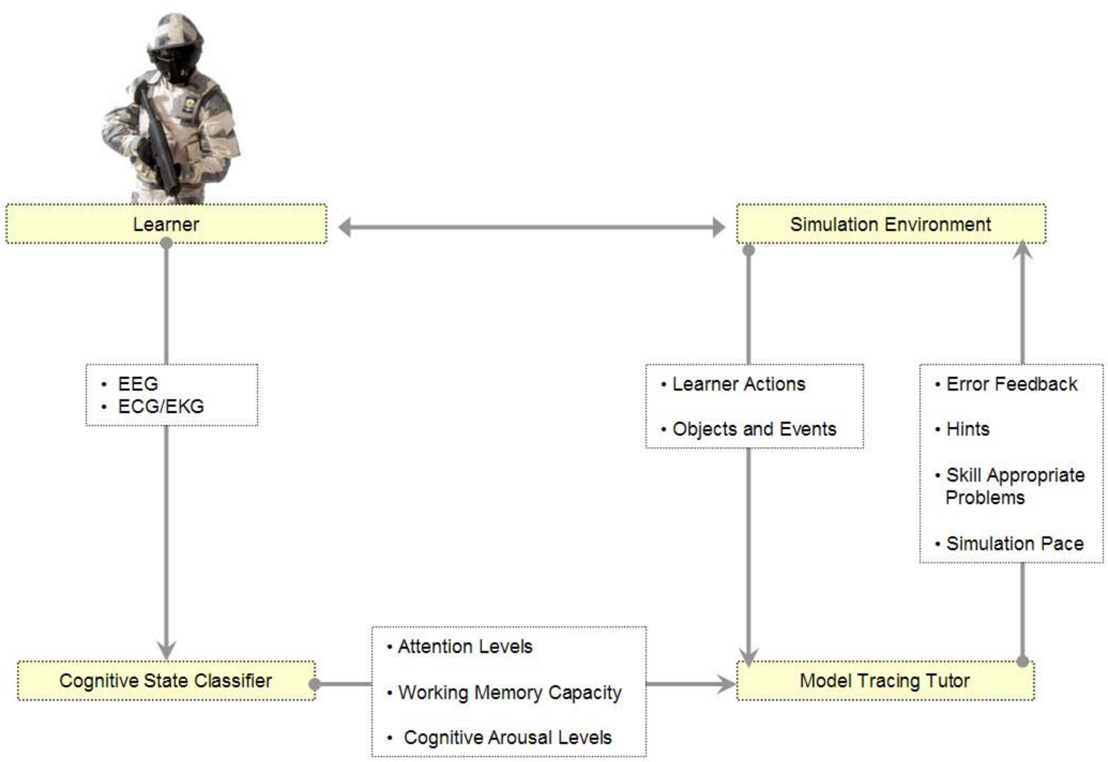

Figure 2: Agmented Tutor Architecture

4 Conclusion

We have described a set of technical solutions — cognitive tutors and neurophysiological sensing — that could combine to improve the efficiency and effectiveness of simulation based training systems. The joint use of these technologies will provide a diagnostic capability of unprecedented scope and will have broad relevance across simulation based training platforms. An instructional system embodying these features could dynamically guide students toward expertise — with feedback based on both an assessment of overt problem solving actions and parameters such as working memory load, attention, and cognitive arousal that have an impact on learning outcomes. These capabilities could combine to transform training simulations from platforms that support unstructured practice to environments that intelligently guide students towards effective and efficient performance.

Acknowledgements

This research was supported by a contract with DARPA and funded through the U.S. Army Natick Soldier Center, under Contract No. DAAD-16-03-C-0054, for which CDR (Sel.) Dylan Schmorrow serves as the Program Manager of the DARPA Augmented Cognition program and Mr. Henry Girolamo is the DARPA Agent. Any opinions, findings, conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Army.

References

Anderson, J. R. (1993). Rules of the mind. L. Erlbaum Associates:Hillsdale, N.J.

Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. The Journal of Learning Sciences, 4, 167-207.

Anderson, J. R., Bothell, D., Byrne M. D., Douglas S., Lebiere, C. & Qin, Y. (in press). An Integrated Theory of the Mind. Psychological Review.

Anderson, J. R., & Gluck, K. (2001). What role do cognitive architectures play in intelligent tutoring systems? In D. Klahr & S. M. Carver (Eds.), Cognition & instruction: twenty-five years of progress (pp. 227–262). Erlbaum.

Anderson, J. R. & Lebiere, C. (1998). The Atomic Components of Thought. Mahwah, NJ: Lawrence Erlbaum Associates.

Bloom, B.S. (1984). The Two Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. Educational Researcher, Vol. 13, Nos. 4-6

Byrne, M. D., Kirlik, A., Fleetwood, M. D., Huss, D. G., Kosorukoff, A., Lin, R., & Fick, C. S. (2004, in press). A closed-loop, ACT-R approach to modeling approach and landing with and without synthetic vision system (SVS) technology. To appear in Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society.

Corbett, A.T. and Bhatnagar, A. (1997). Student modeling in the ACT Programming Tutor: Adjusting a procedural learning model with declarative knowledge. Proceedings of the Sixth International Conference on User Modeling. New York: Springer-Verlag Wein

Gevins A, Smith ME, McEvoy L, Yu D (1997) High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex 7:374–385.

Jung, T-P, Makeig, S, Stensmo, M, and Sejnowski, TJ, (1997) "Estimating alertness from the EEG power spectrum," IEEE Transactions on Biomedical Engineering, 44(1), 60-69

Kautz, H. and Allen, J. (1986). Generalized plan recognition. In Proceedings of AAAI-86, pages 32–37

Kulik J.A.& Kulik C.C. " Computer based instruction: What 200 evaluations say?". Centre for Research and Teaching, University of Michigan, 1988

McBreen, B. B, (2003) Close Combat and learning infantry tactics. Retrieved August 8th 2004 from http://www.tecom.usmc.mil/techdiv/ITK/CCM/CCM.htm

Mayer, R.E. (2001). Multimedia Learning, Cambridge: Cambridge University Press

Salvucci, D. D. (2001). Predicting the effects of in-car interface use on driver performance: An integrated model approach. International Journal of Human-Computer Studies, 55,85-107.

Schooler, J.W., Reichle, E.D., & Halpern, D.V. (in press) Zoning-out during reading: Evidence for dissociations between experience and meta-consciousness. In D.T. Levin (Ed), Thinking and Seeing: Visual Metacognition in Adults and Children. MA: MIT Press.

Sweller, J. (1988). Cognitive load during problem solving: Effects of learning. Cognitive Science, 12, 257-285.

United States Marine Core (2004) USMC Family of tactical decision making simulations. Retrieved August 8th 2004 from http://www.onr.navy.mil/sci_tech/special/353_exped/hptetds.asp

Zachary, W., Cannon-Bowers, J., Bilazarian, P., Krecker, D., Lardieri, P., & Burns, J. (1999). The Advanced Embedded Training System (AETS): An intelligent embedded tutoring system for tactical team training. International Journal of Artificial Intelligence in Education, 10, pp. 257-277.