Intelligent Tutoring: Bridging the Gap from Knowing to Doing

Bruce M. Perrin, Barbara J. Buck, Sara Elizabeth Gehr
The Boeing Company
St. Louis, MO
Bruce.m.perrin@boeing.com, Barbara.j.buck@boeing.com, Liz.gehr@boeing.com

ABSTRACT

Often, our fielded training systems emphasize media for imparting facts, rules, and procedures, followed by simulations or live events where trainees are expected to apply that knowledge. Seemingly missing from this sequence, however, is any support for acquiring the cognitive capabilities that underlie task performance. The result can be trainees who know the facts, but not where and how to use them. Presumably, if we provided specific training on the abilities needed to comprehend, assess, and decide on a course of action, we could facilitate the transition from knowing to doing. We tested this exact hypothesis in a training effectiveness study. We first provided students with interactive multimedia instruction (IMI) on the Marine Air Ground Task Force XXI (MAGTF XXI) simulation controls and displays and the tactics for performing an in-stride breach of a minefield. After a test trial with the MAGTF XXI simulation to establish a performance baseline, half of the participants were randomly assigned to Intelligent Tutoring System (ITS) training. The ITS lesson was designed to help the trainees picture the situation the same way experts did and to select actions that implemented the experts' preferred solution strategies. The control group conducted a self-directed review of the IMI, allowing them to study material related to any problems they had experienced in the simulation. After each of two IMI or ITS study sessions, the participants performed the in-stride breach. Across trials, the ITS-trained group showed significantly accelerated mastery of the skill in the simulation environment. Clearly, using an ITS to provide explicit training on an expert's problem representation and solution strategies can accelerate learning and optimize investments in more costly and complex simulations and live training events.

ABOUT THE AUTHORS

Bruce M. Perrin is a Boeing Technical Fellow, assigned to the Instructional Systems Design (ISD) team within the Training Systems & Services (TSS) Capability Center. He has been employed at Boeing for more than 25 years, during which time he has been responsible for the analysis, design, and development of training and decision support systems and technologies, and for conducting formative and summative evaluations of them. He is also an adjunct faculty member at Washington University in St. Louis, teaching classes in the Applied Psychology of Learning.

Barbara J. Buck has 21 years of human factors experience at Boeing, both in applied research and advanced design applications. She is currently Principal Investigator for the Advanced Instructional Systems Tool Suite within the Instructional Systems Design (ISD) team of Boeing Training Systems & Services. Her areas of specialization include the application of cognitive psychology, human factors, and instructional design principles to research and development, as well as design solutions.

Sara Elizabeth Gehr has seven years of training program evaluation and human factors experience at Boeing, emphasizing studies of team and individual performance in international coalition distributed mission operations. She is currently assigned to the Instructional Systems Design (ISD) team within Training Systems & Services, where she leads the human factors effort for the development and evaluation of Intelligent Tutoring System authoring and delivery systems.

Intelligent Tutoring: Bridging the Gap from Knowing to Doing

Bruce M. Perrin, Barbara J. Buck, Sara Elizabeth Gehr
The Boeing Company
St. Louis, MO
Bruce.m.perrin@boeing.com, Barbara.j.buck@boeing.com, Liz.gehr@boeing.com

Over the last 45 years, there has been a substantial amount of research on the differences between experts and novices in their job task and problem-solving performance. In domains as diverse as chess (de Groot, 1965) to computer programming (Jeffries, Turner, Polson, & Atwood, 1981) to maintenance troubleshooting (Gitomer, 1984; Glaser et al., 1985), researchers have identified and documented systematic, expert-novice differences in how problems are represented and how they are solved. While there is substantial variation in the terminology used across this diverse set of domains, there appear to be several consistent findings (Sternberg, 2003; Anderson, 2004; Clark et al., 2007). Chief among the findings are differences in the ways problems are represented, with experts using representations that facilitate a solution; differences in the application of overarching strategies that guide the selection of actions; and differences in the degree to which sets of tactical actions have become "automated" through repeated use.

Presumably, providing explicit training on job task and problem-solving abilities such as these would facilitate skill acquisition. In response, cognitive task analysis (CTA) procedures have been developed specifically to expose the knowledge structures and strategic and tactical problem-solving processes of experts, so that they can be used in instruction. Using CTA methods, Means and Gott (1988), for example, speculated that 50 hours of training on experts' representations of the equipment and their maintenance troubleshooting strategies provided the equivalent of five years of job experience. The experimental data on this hypothesis, however, is quite limited, although generally positive. Lee (2004), for example, conducted a meta-analysis study that used literature searches of 10 major academic databases covering the 19-year period of 1985 to 2003. Only seven studies were found that provided experimental data on the effect of CTA-based training on task performance.

Our objective in conducting this work was to add to this research base. Specifically, we studied a simple training sequence of Interactive Multimedia Instruction (IMI) followed by simulation-based training. The IMI was used to present and test understanding of the declarative knowledge of the task – the facts, rules, and procedures used in performing it. Students then practiced the use of this knowledge in simulation-based training scenarios. The primary question under study was whether the addition of explicit abilities training, in the form of an Intelligent Tutoring System (ITS) lesson on an expert's strategies and ways of representing the problem would facilitate learning to perform in the simulation.

In general terms, ITSs operate by comparing a student and an expert model. The expert model is designed to capture the ways an expert represents and/or solves problems in the domain and may be built using results from a CTA. The student model is formed and dynamically adjusted as a student solves problems that the ITS poses. These models are compared, and a third ITS model, an instructional or pedagogical model, then mediates how the estimated differences are translated into instructional interventions. Beyond this somewhat standard, three-model architecture, however, there is little consistency in how ITSs are constructed, how the comparisons are performed, or what actions are taken when differences are identified. A number of ancillary capabilities are also often included in an ITS, e.g., mixed initiative dialog, system learning, dynamically generated explanations or feedback, etc. Assessing ITSs, as a result, is not only about measuring the impact of the system on learning; it should also involve a description of the features and capabilities of the system that could be responsible for the effect (Gold, 1998). We address that objective in the next section.

The Intelligent Tutoring System Implementation

Figure 1 overviews our implementation of the ITS. Our student model implements a profile of dynamically maintained variables, with each variable corresponding to one learning objective. These variables are

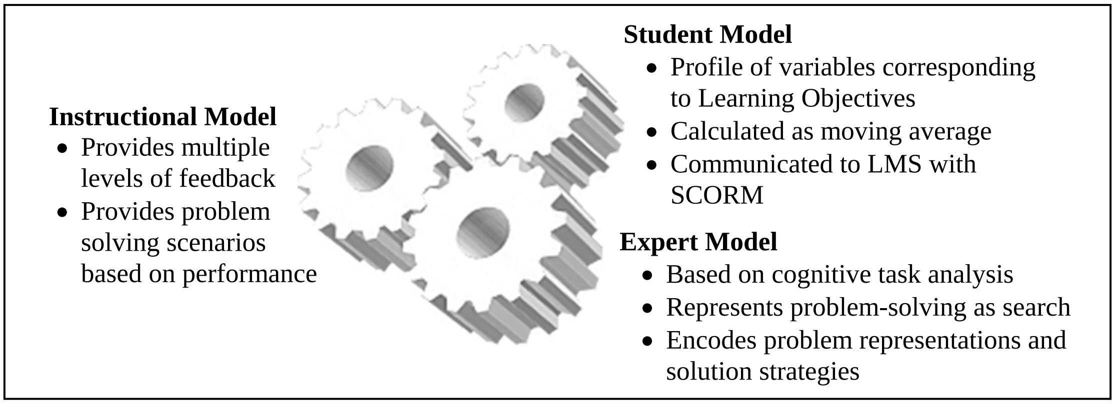

Figure 1. Specific Model Implementations in ITS

estimated as moving averages over a specified number of observations. As a result, changes due to learning are reflected across exercises, as the average increases due to correct performance, or decreases as errors are made. The amount that scores are changed can be adjusted according to the degree to which the action reflects mastery of the learning objective. Amount of change is also adjusted according to the degree of support the ITS provided to the student in selecting this action.

With student requests for assistance or student errors, our instructional model responds with information on problem-solving strategies or problem representation. The specificity of the information increases as additional requests are made or additional errors occur. The instructional model is also tasked with selecting follow-on exercises. By examining current student model scores, the instructional model identifies exercises that represent the next step in skill development or that are needed to clarify diagnostic information (see Perrin, Buck, Dargue, Biddle, Stull & Armstrong, 2007 for additional detail).

Our implementation of the expert model is based on the CTA technique known as PARI, for Precursor, Action, Results, and Interpretation (Hall, Gott, & Pokorny, 1995). Although PARI was developed for the analysis of maintenance troubleshooting tasks, it is based on a more general view of problem-solving as search through a problem space (Newell & Simon, 1972). PARI includes standard procedures that can be used to identify representative problem sets for ITS exercises. It also provides methods to elicit detailed information from experts on how they represent a given state of a solution (what issues have been resolved and what issues remain), optimal and alternative paths to a solution, and their strategies for selecting actions at each step along those paths. Our expert model directly encodes these solution paths. For each path, the model also encodes the expert's summary of the situation (representation of the problem) and the rationales for the possible next steps (see Figure 2). Additional detail on our ITS architecture and implementation can be found in Perrin (2009).

THE TRAINING EFFECTIVENESS STUDY

To test our primary hypothesis that specific abilities training would facilitate the development of problem-solving skills during advanced, follow-on training, we needed to identify the following:

- A cognitively complex problem-solving skill that would benefit from abilities training. Simply put, there would be little to gain from training abilities, if there are few cognitive demands in performing the skill; and
- A simulation, live, or other environment where skill acquisition could be measured following abilities training.

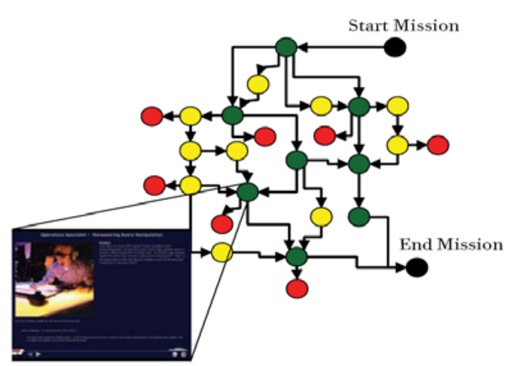

Figure 2. Optimal and Alternative Decision Paths in the Expert Model

We met both of these objectives by using a prototype system we had developed under contract for the Joint Advanced Distributed Learning Co-Laboratory (Joint ADL Co-Lab) in Orlando, FL (Biddle, Perrin, Dargue, Pike, & Marvin, 2006).

The skill taught by the prototype is that of coordinating and directing an in-stride breach of a minefield as the commander of a mechanized infantry – tank team. Although the in-stride breaching tactic is relatively well defined, it meets the requirement for cognitive complexity. It requires initial planning and continuing, time-critical problem solving as the scenario unfolds. Abilities to assess situations, to evaluate alternatives, and to coordinate actions are critical. Formative studies conducted as part of this project supported the premise that the skill was cognitively complex. High rates of error, particularly errors of omission, were common initially, as the problem-solving skill built through practice.

In the prototype, the skill acquisition environment is provided by the Marine Air Ground Task Force XXI (MAGTF XXI) simulation. MAGTF XXI is a real-time, High Level Architecture (HLA) conformant, tactical simulation built for the U.S. Marine Corps. It was developed by MAK Technologies under the Program Manager Training Systems (PM TRASYS) Tactical Decision-making Simulation (TDS) program to facilitate expeditionary warfare training. Measuring skill acquisition was also facilitated, as the prototype implemented automated performance assessment algorithms that we had designed and built for the project. For the scenarios used in this study, the prototype could detect and evaluate 21 distinct actions.

Use of the automated performance assessment capabilities of the prototype was also noteworthy, as it eliminates experimenter bias as a possible threat to the internal validity of the study. Simply put, experimenter expectations could not influence performance data that were automatically generated by the prototype. Additional detail on this prototype and the simulation task are provided in Biddle, Perrin, Dargue, Pike, and Marvin (2006).

METHOD

Participants

Twenty-four Boeing employees volunteered for the study. The group was composed of 20 males and 4 females. Six had prior military experience, but individuals with previous experience with infantry-tank task force operations were disqualified from participation.

Procedure

Participants first read and signed an Informed Consent form. It described their rights to anonymity and to withdraw at any time, as well as the general features of the study. Next, they completed a brief demographic questionnaire. It included questions on age, gender, education, and the frequency of computer games use.

All participants then received Interactive Multimedia Instruction (IMI) that covered the in-stride breaching tactics and the MAGTF XXI interface. The instruction included descriptions, photographs, and graphics of the equipment; annotated screen shots and narrated video clips of the MAGTF simulation displays and commands; and annotated screen shots illustrating breaching operations and tactics. The participants had 30 minutes to complete these materials, and all of them did so, with a few moments for review. Following study of the IMI, all participants were given a knowledge test. The knowledge test consisted of 63 multiple-choice or matching questions. The tests were scored and returned to the participants with corrections, so they could review their performance and improve their understanding of the breaching tactics and simulation displays and controls.

Next, each participant used the MAGTF XXI simulation to conduct the initial phases of an in-stride breach. Each scenario ended when the units were positioned for the breach and when fire to suppress the opposing force and mask your units' movements was established. The scenarios were limited to these steps so that they would be short enough to allow multiple trials. Participants were given 10 minutes to complete these actions. To this point, all participants had received the same instructions and training; this initial trial established a performance baseline. After it, the participants were randomly assigned to one of two training treatments as follows.

IMI Review (Control)

Under the control condition, participants were given 20 minutes to conduct a self-directed review of the IMI and take and receive feedback on a knowledge test. The IMI review allowed them to focus on any aspects of the breaching tactics or the simulation controls and displays that they wished, based on problems they had experienced during the baseline trial. This review was followed with another test, which was scored and returned to provide feedback, as before. This test had the same questions as the first, but in a different order. Because reviewing the IMI and receiving feedback from a knowledge test represents the control condition of this study, it is important to note that this IMI was originally created to provide training for the prototype system. It represents an independent product, designed with the objective of fully preparing users for the prototype system, and not merely as a control condition for this study.

Following the self-directed review and test, participants in the IMI Control group conducted a second in-stride breach using MAGTF XXI. This trial was followed with a final round of self-directed IMI review, knowledge testing with feedback, and an in-stride breach using MAGTF XXI. Thus, overall, each participant conducted three breaching operations, the first providing a baseline followed by two test trials. Each scenario was identical, so improvement in performance would represent a combination of study of the declarative knowledge in the IMI, knowledge testing feedback, and practice on the simulation.

Intelligent Tutoring System (Experimental)

Following the baseline trial, participants in this group studied an ITS lesson built using the architecture described previously. The ITS lesson did not repeat the declarative knowledge covered in the IMI. Rather, it presented a mission that required an in-stride breach that was similar to, but different from one used in the test trials. It required that the participants select an action to implement along the optimal or alternative solution paths. It provided hints on the next actions, based on the expert's preferred solution strategy, when the student asked for assistance. It also provided feedback on the student's actions, including the impact on problem resolution and the optimal action, if it was different from what the student had selected. The ITS required the students to apply their knowledge of the MAGTF XXI interface and breaching tactics. For example, the IMI described how to use light vegetation and terrain features to help mask the movement of units. It also provided screen shots and interpretations of terrain displays. The ITS lesson, on the other hand, reached points in problem solutions where a unit needed to be moved, and asked the student to select a route. That selection, of course, depends upon the student using his/her knowledge of terrain masking (see Figure 3).

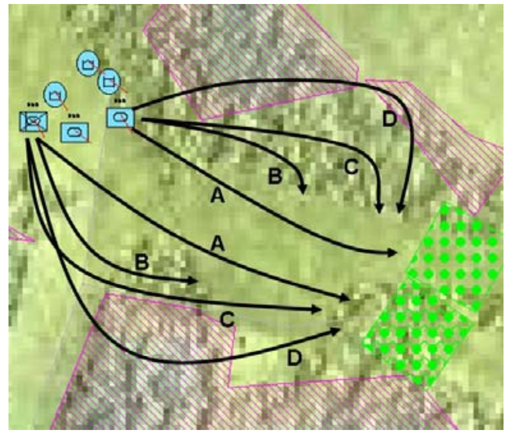

Figure 3. Selecting Routes in the ITS

After completing the ITS lesson, this group conducted a second in-stride breach using MAGTF XXI. This trial was followed with a final round of ITS study and a final in-stride breach. Similar to the control participants, each participant in the ITS group conducted three breaching operations, with one providing a baseline and two test trials.

The average time students took to study the ITS lesson was about 20 minutes the first time and about 15 minutes the second time. These study times were approximately the same as the time the control participants spent reviewing the IMI and their knowledge test results. Overall, the study times between the groups were approximately equal.

Data and Analysis

The primary data used for this study were the 21 actions that were automatically monitored and scored during the MAGTF XXI in-stride breach test scenarios. To help organize these data and provide more stability in the measures, the actions were grouped into one of six behaviors, as follows:

1. Establish Support by Fire: Assign and position units to provide support by fire.
2. Coordinate Movement: Coordinate arrival of support by fire units.
3. Position for Breach: Assign and position other units to perform or support the breach.
4. Identify safe route: Establish route that minimizes exposure to hostile fire.
5. Control Suppressing Fire: Establish locations and frequency for suppressing fire.
6. Control Obscuring Fire: Establish locations and frequency of fire to help mask your movements.

Scores on these behaviors were the proportions of correct actions to the total number of opportunities. These six measures were used in a treatment by trials, repeated measures analysis of variance (ANOVA).

RESULTS

We conducted several tests for pre-existing differences between our two treatment groups. Two-group t-tests were run on several of the demographic factors, and no significant differences between the groups were identified: age (**t**(21) = -0.58, **p** < .568); years of education (**t**(21) = 0.84, **p** < = 0.408); and frequency of playing computer games (**t**(21) = -0.63, **p** < 0.535). We also examined differences between our treatment groups in their initial knowledge of breaching tactics and the MAGTF interface using the first knowledge test. The difference in this initial knowledge test score did not reach statistical significance (**t**(21) = 0.43, **p** < 0.672). Based on these results, there is little reason to believe there were any substantial, pre-existing differences between our two treatment groups.

For the primary test of our hypothesis, the results from the repeated measures ANOVA are provided in Table 1. For three of the six behaviors, the trial by treatment interaction was significant, indicating different patterns of skill acquisition in the simulation depending on whether the participants reviewed the IMI or studied the ITS. Each of these three behaviors also showed either a trial or a treatment main effect. For the other three behaviors, all treatment, trial, and interaction effects did not reach significance, except for one trial main effect.

To explore the nature of these results, we averaged the scores of the three behaviors that were affected by the training treatment (e.g., the first three rows of Table 1). We also averaged the three behaviors that were not influenced by the training treatment. These calculations produce two composite measures that we used for descriptive purposes only – one composed of nine actions representing three behaviors that were apparently influenced by the training treatment, and one composed of 12 actions representing three behaviors that were not affected by the treatment. In Figure 4, scores on these two composite measures are shown across trials for the IMI and ITS training treatments.

The difference seems apparent. For those behaviors affected by the training treatment, the initial performance of both groups was relatively poor. Both groups exhibited appropriate behaviors less than 20% of the time. Over the course of studying the ITS lesson, however, performance of the experimental group improved steadily. This group started showing successful performance in only about one in five opportunities (.21) in Trial 1 (baseline) and ended the training with performance approaching 60% (.57). The Control group, on the other hand, showed virtually no change in performance, and if anything, showed a slight downward trend from .28 in Trial 1 (baseline) to .24 in Trial 3.

Table 1. Summary of Significant Effects from ANOVA

Simulation Behavior	Treatment	Trial	Trials X Treatment
Establish Support by Fire	-	F(2,44) = 21.88; p < 0.001	F(2,44) = 3.14; p < 0.05
Coordinate Movement	-	F(2,44) = 6.25; p < 0.004	F(2,44) = 3.72: p < 0.032
Identify safe route	F(1,22) = 6.66; p < 0.017	-	F(2,44) = 5.45; p < 0.008
Position for Breach	-	F(2,44) = 4.51; p < 0.017	-
Control Suppressing Fire	-	-	-
Control Obscuring Fire	-	-	-

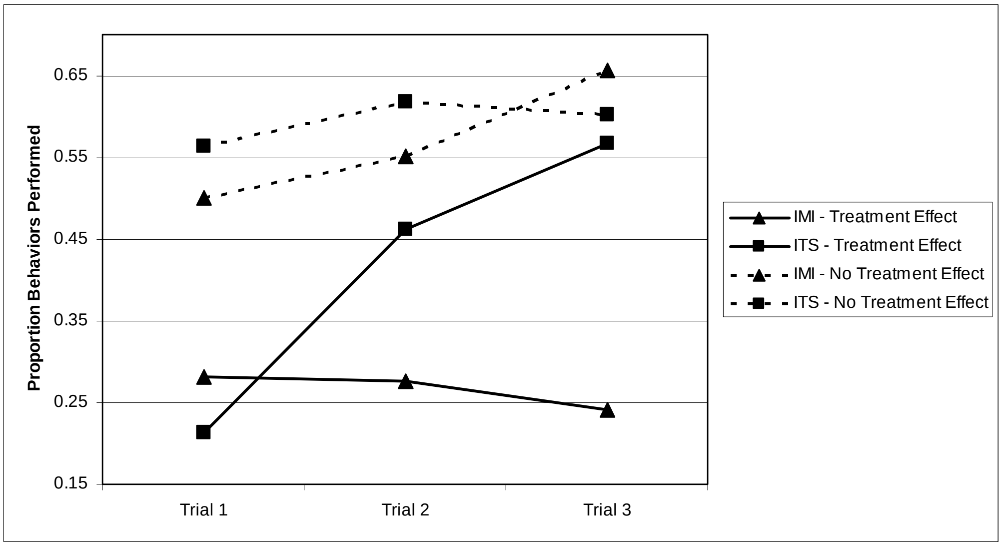

Figure 4. Learning Curves for Behaviors Affected or Unaffected by Treatment

For the behaviors showing no treatment effects (no treatment main effect and no treatment by trial interaction), both groups showed initially high performance that changed little during the training. The IMI Control group initially performed 50% of the actions successfully, and finished at 65%. The comparable performance scores for the ITS Experimental group were 56% and 60%.

DISCUSSION AND CONCLUSIONS

Clearly, providing explicit training on abilities accelerated learning in the follow-on simulation scenarios. In fact, it appears that it did more than accelerate skill acquisition; it enabled learning for a large portion of the skill. Students who were limited to studying and getting feedback on the IMI and observing their performance in the simulation were making little or no progress on mastering a substantial subset of the in-stride breach behaviors (42.8% of the actions that we measured). Some explicit intervention, clarifying how, why, or when to apply declarative knowledge appears to be crucial to attaining substantial aspects of this skill.

Post analysis, there seems to be a consistent difference between behaviors that benefited from explicit abilities training and those that did not. The behaviors dealing with controlling fire and positioning for breach, which did not benefit from ITS training, seem relatively straightforward. For example, the IMI on controlling fire specified the type, location, and pace of firing, with few, if any, requirements for evaluation and decision making. The transparency of these behaviors is also supported by the initial high rate of performance. Participants in both groups performed well on these behaviors with only the initial IMI training and performance changed little with additional training of either type. The difficulty in implementing these behaviors seemed more a matter of remembering them, with all the other on-going demands on attention, rather than deciding when or how to use them.

The behaviors that were impacted by explicit abilities training, on the other hand, seemed to include a component of evaluation and judgment. Consider the behavior of identifying a safe route for transit. The principles involved appear simple enough, e.g., during movement, use terrain elevations and light vegetation to help conceal your assets from suspected areas of opposing forces. The IMI contained seven screens covering the topics of reading topographical maps and using terrain and vegetation to mask movement. The knowledge test, which was used to provide feedback on learning, contained six questions relevant to these topics. As the IMI Review group scored quite high on the final test (90.8% correct), it seems clear that these participants knew this information. Additionally, participants in the IMI Control group would have observed their units taking fire from the opposing forces during transit, which should have caused them to focus on these principles. Even so, they made no apparent progress in mastering the use of this tactic during the simulation scenarios.

The ITS lesson did not further review the principles related to safe transit. Those were known. Rather, in a similar, but distinct situation, the ITS required the student to implement these principles when and where they were needed. A situation requiring unit movement occurred naturally three times in the ITS training scenario. In those situations, students were given different paths, similar to what was shown in Figure 3. Students selected one action. For a valid selection, the ITS reinforced the decision, reviewing the elements of the situation that drove this choice from the expert's perspective, e.g., from the perspective of how the problem can be represented. It also reviewed the strategies and tactical actions that were the basis for this action. For incorrect responses, the ITS provided feedback based again on problem representation and strategy and tactics implemented, if any. The feedback covered the aspects of the problem unresolved, or perhaps even worsened, due to the action taken. Generally, the correct response was not given at this point; rather, the student was returned to the decision point to reconsider the remaining options. The impact of a valid response on problem resolution and strategy and tactic use was reviewed after the student took that path.

Given the impact of explicit abilities training on skill acquisition, the question becomes one of whether the impact is of practical significance. In other words, is the cost of developing and implementing explicit abilities training in an ITS offset by the reduction in time or staffing it can produce? A full scale training utility analysis is beyond the scope of this paper, but an informal analysis suggests that the answer is yes, even when relatively small numbers of students are to be trained.

Since participants without abilities training do not appear to be making substantial progress in mastering some portions of the skill, we can assume that we will need to add some explicit instruction. The appropriate comparison appears to involve the two training approaches below; each includes time requirements, either based on our study or estimated.

1. IMI, then ITS (35 minutes), then Simulation (30 minutes). This approach replicates the experimental treatment in this study. It includes the time for the first simulation exercise, even though it was designed only to provide a baseline of performance, as this trial might have affected learning.
2. IMI, then Simulation (estimated at 30 minutes), then Instructor Debrief (estimated at 42 minutes). This sequence is a modification of the control treatment that adds instructor-led training on abilities.

One might ask: Why not add abilities training to the IMI? In fact, we consider the implementation of ITS that we used in this study as type of IMI. The ITS is basically Level 3 IMI that features high levels of interactivity; extensive branching that is based on comparing a student's problem-solving actions to those of an expert; and feedback based on the problem representation and problem-solving strategies and tactics of an expert.

Comparing the two approaches above, IMI is a constant. So, we can focus on the tradeoff between ITS development and delivery costs vs. the delivery costs of instructor-led simulation feedback for abilities training. For ITS development, our lessons draw much of their data from standard training products, e.g., learning objectives, standard exercise datasets, etc. (Perrin, 2009). As a result, ITS lessons can be constructed rapidly. The lesson used in this study, for example, was built in a little over seven hours. It should be noted, however, that this development benefited from the re-use of existing graphics and was built by an individual very familiar with the training task and our ITS authoring templates and methods. To provide a more conservative estimate of the development effort, we will use 25 hours.

To estimate delivery costs, we will need to estimate student and instructor time requirements under the second approach. If we assume that instructor-led debriefs are similar to instructor-led lectures, in terms of time requirements, there is an extensive research literature indicating that IMI reduces student contact time by about 20-30% compared to instructor-led training (Johnston, 1995; Kulik, 1994). If this relationship holds, then we could expect the instructor-led debriefs to require 42 minutes (20% more than the ITS) for abilities training. We will exclude instructor time to observe the students during the simulation exercises, under the assumption that they would be present in any case. If they were present solely to provide abilities training, instructor time requirements should be increased by an additional 20 minutes in order to observe two, 10-minute simulation scenarios. Students, on the other hand, need to complete three, 10-minute simulation exercises, in order to complete the same amount of practice following abilities training as approach #1. The overall student time requirements would be 72 minutes.

Under these assumptions, explicit abilities training using the ITS on this single topic should reduce student time by 7 minutes and instructor time by 42 minutes.

With as few as 36 students, the savings of instructor time would exceed the development time for the ITS. Total student throughput time would have also been reduced by 4.2 hours (7 minutes times 36 students). If we include the time to observe two simulation scenarios as part of the instructor's time commitment, the breakeven point is only 24 students. When the simulation environment is particularly expensive to operate, including high-end devices or live events, even smaller numbers of students would be required to justify the development of an ITS.

There is also reason to believe that our estimate of the instructor's time requirement for abilities training during debrief may be conservative. A growing body of research suggests that providing this type of training may not be as simple as asking instructors to discuss how they solve problems, even if the instructor is an expert. When comparing experts' description of how they solve problems to specific solutions they have produced, experts often unintentionally omit steps (Johnson, 1983; Collins, Green, & Draper, 1985; Chao & Salvendy, 1994), cite factors in their descriptions that have no apparent relationship to the solutions they produced (Einhorn, 1974; Cooke & Breedin, 1994; Kareken & Williams, 1994), or overlook factors that are associated with their actions (Crandall & Getchell-Reiter, 1993). An ITS or other method based on a thorough CTA may be even more efficient in imparting an expert's abilities than our informal analysis suggests.

A final approach to proving explicit abilities training, not discussed above, would be to use automated methods for performance assessment and then, provide abilities feedback in the simulation. This capability could support the instructor, and thus reduce the staffing demands for covering abilities. That option, essentially, places the ITS capability in the simulation environment, rather than as a type of IMI. Such an approach has the advantage of using more realistic behaviors in the diagnosis and feedback cycle, but at the cost of developing methods that can accurately monitor and assess behavior in a complex, dynamic simulation.

Our ITS architecture was designed, in part, to enable its use as part of either IMI content or simulations. That is, rather than screen shots with discrete alternatives, as shown in Figure 3, the ITS would be integrated with the simulation and evaluate the student's commanded routes. The prototype system used in this study was employed solely to provide automatically scored behaviors. It did, in fact, implement our ITS student model as part of a study on learning needs diagnosis (Perrin, Buck, Dargue, Biddle, Stull & Armstrong, 2007). The data needed to evaluate a simulation-based ITS, compared to the alternatives above, is beyond the scope of this paper, but the capability exists.

The way experts see their task and select their courses of action are different from those of novices. This seems clear. What we have provided is empirical data, consistent with a small but growing body of research, which indicates that providing instruction via an ITS that depicts the experts' representations and solution strategies can facilitate skill learning. In short, this type of training can help to bridge the gap between knowing and doing.

REFERENCES

Anderson, J.R. (2004). *Cognitive psychology and its implications*. New York, NY: Worth Publishers.

Biddle, E., Perrin, B., Dargue, B., Pike, W.Y., & Marvin, D. (2006, December). *Performance based advancement using SCORM 2004.* Paper presented at the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

de Groot, A.D. (1965). *Thought and choice in chess*. The Hague: Mouton.

Chao, C.J., & Salvendy, G. (1994). Percentage of procedural knowledge acquired as a function of the number of experts from whom knowledge is acquired for diagnosis, debugging and interpretation tasks. *International Journal of Human-Computer Interaction*, *6*, 221-233.

Clark, R.E., Feldon, D.F., van Merrienboer, J., Yates, K.A., & Early, S. (2007). Cognitive task analysis. In Spector, J.M., Merrill, M.D., van Merrienboer, J., & Driscoll, M.P. (Eds.), *Handbook of Research on Educational Communications and Technology*, New York, NY: Routledge.

Collins, H. M., Green, R. H., & Draper, R. C. (1985). Where's the expertise?: Expert systems as a medium of knowledge transfer. In M. Merry (Ed.), *Proceedings of the fifth technical conference of the British Computer Society Specialist Group on Expert Systems '85* (pp. 323-334). New York: Cambridge University Press.

Cooke, N. J., & Breedin, S. D. (1994). Constructing naive theories of motion on-the-fly. *Memory and Cognition*, *22*, 474-493.

Crandall, B., & Gretchell-Leiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the "intuition" of NICU nurses. *Advances in Nursing Science, 16*(1), 42-51.

Einhorn, H. (1974). Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology, 59,* 562-571.

Gitomer, D.H. (1984). A cognitive analysis of a complex troubleshooting task. Unpublished doctoral dissertation, Pittsburgh, PA: University of Pittsburgh.

Glaser, R., Lesgold, A., Lajoie, S., Eastman, R., Greenberg, L., Logan, D., Magone, M., Weiner, A., Wolf, R., Yengo, L. (1985). Cognitive task analysis to enhance technical skills training and assessment. (Final Report to the Air Force Human Resources Laboratory on Contract No. F41689-8v3-C-0029.) Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh.

Gold, S.C. (1998). The design of an ITS-based business simulation: A new epistemology for learning. *Simulation and Gaming, 29,* 462-474.

Hall, E.M., Gott, S.P., & Pokorny, R.A. (1995). *A procedural guide to cognitive task analysis: The PARI methodology*. Technical Report No. AL/HR-TR-1955-0108. Brooks AFB, TX: AFMC.

Jeffries, R.P., Turner, A.A., Polson, P.G., & Atwood, M.E. (1981). The processes involved in designing software. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Johnson, P. E. (1983). What kind of expert should a system be? *The Journal of Medicine and Philosophy, 8,* 77-97.

Johnston, R. (1995, June). The effectiveness of instructional technology: A review of the research. *Proceedings of the Virtual Reality in Medicine and Developers' Exposition*. Cambridge, MA: Virtual Reality Solutions, Inc.

Kareken, D. A., & Williams, J. M. (1994). Human judgment and estimation of premorbid intellectual function. *Psychological Assessment, 6*(2), 83-91.

Kulik, J.A. (1994). Meta-Analytic Studies of Findings on Computer-Based Instruction. In Baker, E.L. and O'Neil, H.F. (Eds.). Technology Assessment in Education and Training. Hillsdale, NJ: LEA Publishers.

Lee, R. L. (2004). *The impact of cognitive task analysis on performance: A meta analysis of comparative studies*. Unpublished Ed.D. dissertation, Rossier School of Education, University of Southern California, USA.

Means, B. & Gott, S. (1988) Cognitive task analysis as a basis for tutor development: Articulating abstract knowledge representations. In J. Psotka, L. D. Massey, & S. A. Mutter, Intelligent tutoring systems: Lessons learned. Hillsdale, N.J. : Lawrence Earlbaum.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Perrin, B. (2009). *Intelligent Tutoring Systems: Facilitating Learning While Holding to Standard Practice*. Paper presented at the International Training and Education Conference, Brussels, Belgium.

Perrin, B., Buck, B., Dargue, B., Biddle, E., Stull, T., & Armstrong, C. (2007, December). *Automated Scenario-Based Training Management: Exploring the Possibilities.* Paper presented at the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Sternberg, R.J. (2003). *Cognitive psychology*. Belmont, CA: Wadsworth/Thomson Learning.