AUTOMATED ADAPTIVE INSTRUCTION FOR EMBEDDED TRAINING

Kent E. Williams
University of Central Florida
Institute for Simulation and Training
Orlando, Florida

Richard E. Reynolds
Naval Training Systems Center
Orlando, Florida

Thomas F. Carolan
University of Central Florida
Institute for Simulation and Training
Orlando, Florida

ABSTRACT

Empirical evidence from cognitive learning research suggests guidelines for structuring instructional information in a manner which is consistent with the way people process it. A cognitive engineering process based on these guidelines was used to restructure the content of an existing embedded training lesson. A heuristic was developed to select and adaptively sequence the most appropriate exercises for the individual trainee based upon the individual's history of performance. An experimental evaluation of the effectiveness of the cognitively engineered instructional content and the adaptive sequencing strategy was conducted. Trainees using the cognitively engineered lesson made 65 percent fewer errors on a performance test than the control group. The results also demonstrate that adaptive exercise sequencing will increase both the effectiveness and efficiency of computer based training. These results are discussed in the context of the application of intelligent tutoring system research to embedded training.

INTRODUCTION

A major component of Fleet readiness involves the skill level of the personnel who are responsible for interacting with sophisticated sensor and weapon system consoles. Frequent operator training is required to limit perishability of skills and maintain criterion levels of performance.

The initial training process, conducted in classroom environments, leads to the acquisition of job-specific knowledge and the development of basic skills. However, the individual's ability to apply that knowledge and refine those skills requires considerable practice and experience. Without continued refresher training, skills degrade rapidly [7]. Consolidation and refinement of skills is achieved through practice and experience. However, this requires that a training system be available to personnel, and that personnel be periodically removed from duty platforms to participate in a training program.

The concept of embedded training (ET) has been developed as a potential solution to the problem of operator training. By building capabilities into, or adding them onto, an operational system, operator skill proficiency can be maintained and enhanced in an accessible, high fidelity training environment. ET systems provide a simulation or stimulation environment with which the individual can interact to practice needed skills.

An analysis of the advantages and disadvantages of ET, as compared to classroom training, was conducted by Williams and Reynolds [8]. They found that a critical advantage of classroom training is lost when training is transferred to an ET environment. Only a few of the systems referred to as embedded trainers have any built-in instructional technology. The advantage of classroom training is an instructional methodology that promotes efficient acquisition and retention of skills and knowledge. The effective use of ET also requires an instructional technology that promotes learning and retention of critical skills.

Considerable research has been directed at understanding the cognitive processes involved in knowledge acquisition and skill development. This effort has led to the formulation of various learning principles and the generation of methods to assist or enhance learning. Cognitive science research has led to the formalization of many learning processes and principles. When processes and principles can be made formal or explicit, they can be implemented in software as artificial intelligence (AI) programs and architectures.

In ET environments, the skills and knowledge to be learned are very task- or objective-specific, and their application has a specific outcome on the environment. Training is goal-directed; knowledge and skills to be learned and applied are logically associated with specific goals or objectives. It is important to take advantage of this logical association in the training environment.

When training is translated to an ET environment, the instructional dimension is reduced. Williams and Reynolds [8] proposed that this problem can be overcome by applying appropriate software implementations of cognitive science and AI methods and techniques. Sound principles of learning and intelligent tutoring strategies can be implemented as instructional features under control of the software. Many instructional activities, such as adaptive exercise selection and performance feedback, can be automated.

INSTRUCTIONAL CONTENT

Much of the empirical evidence from cognitive learning research suggests that a specification of the content of the individual exercises in terms of explicit rules will improve learning effectiveness [9]. Two assumptions underlie this explicit rule formulation approach to instructional development. The first is that learning to perform a task, such as operating a tactical console, involves developing a cognitive model that represents the essential information and operational procedures required for task performance. This model is then used to guide subsequent interaction with the device. The second assumption is that the underlying architecture of cognition can essentially be defined within a production system framework [1, 2]. From the perspective of a production system architecture, the individual's cognitive model is structured as, and essentially equivalent to, a production system model of the instructional information. The goal of the training system is to guide the development of a cognitive model, or what Anderson, Conrad, and Corbett [3] have referred to as an ideal student model, which provides an accurate representation of the tasks to be performed and the information necessary for their performance. This development is best accomplished by conducting a cognitive analysis to develop a production system model from the instructional material [4, 5]. The result of this cognitive task analysis process is the hierarchical structuring of the task and subtask to be accomplished along with the production system model of the rules which explicitly specify the means for accomplishing these tasks.

Network of Exercises

A large network of exercises is developed from this hierarchical production system. An exercise is associated with each rule and declarative fact in the hierarchy. Each exercise represents an underlying rule that combines relevant facts and/or lower level rules. Trainees are presented with these exercises, and the process of acquiring knowledge continues until all facts and rules are learned. The trainee works through this network to acquire the expertise. When all rules are learned, the student has a complete cognitive model of console operations to be used for interacting with the console. Critical to this implementation is the development of the production system, or set of rules which simulates system functions and operations.

In working through the exercise set for the training system which teaches tactical console operation, the student learns all the declarative facts and how to compose or compile these facts so that when certain cues or subgoals are set, certain actions or sequences of actions are triggered. Each exercise explicitly describes the subgoal or goal state associated with the production represented by the exercise. The exercises are tightly related to the rules and facts they represent. In this way, intelligent processes can deliver exercise content most efficiently and can diagnose any errors in the trainee's development of the cognitive simulation model. This diagnosis looks at which rules have not been learned properly as well as what pieces of declarative knowledge (i.e., facts or concepts) have not yet been properly assimilated or chunked in the formulation of rules. This is an extremely powerful method for diagnosing the strengths and weaknesses of the student learning system operations. This diagnosis helps determine what lessor content should be presented to the student.

Practice & Diagnostics

Each exercise includes an exposition of the information represented by the rule, an example that requires a response from the student to test the student's ability to use the knowledge addressed by the exposition, and a diagnostic that determines what part or parts of the rule (represented by the exposition and its example) have not yet been mastered. A typical sequence for any production to be learned by the student consists of presenting the expository information in the exercise, then presenting a problem or example requiring the application of the production, such as a response to a question or the performance of a procedure. If the behavior is incorrect or inappropriate, the student is presented with a set of diagnostics to isolate the source of the error(s) in terms of a specific rule or fact. Isolating the source of errors to a specific rule enables the trainee to overcome difficulties in acquiring a working cognitive model of console operations.

The student is diagnosed to determine the conditions or actions of a production for which the student's knowledge is weak, as well as that for which the student's knowledge is strong. Since each exercise created employing this methodology can be explicitly linked to pieces of a production in the form of conditions, actions, or other rules, the training program can query the trainee who fails an exercise to determine what conditions, action, or other rules have been or not been learned. Each exercise has one diagnostic that is directly associated with the content of that exercise, as well as additional diagnostics for each rule and declarative fact from which the exercise is composed. Upon failure or success on an exercise, the system searches for and selects an exercise for presentation which overlaps most with what the student knows and least with what the student does not know. This method exploits strong partial knowledge to build new knowledge.

INSTRUCTIONAL STRATEGY--ADAPTIVE EXERCISE SEQUENCING

The instructional strategy employed involves adaptive exercise sequencing, capitalizing on the trainee's strong prior knowledge. Sequencing of exercises plays an important part in the process of new rule assimilation and composition [6]. The capability to select an exercise which overlaps most with what the student knows is built into the instructional system through a selection heuristic. The heuristic selects the exercise with the greatest overlap in partial knowledge with the exercises the student has already mastered and the least overlap with new or previously failed elements.

Exercise Selection Heuristic

To keep track of the student's progress through the lesson, the system keeps a record called the student model. The student model records the status of each exercise as the lesson progresses. The status of each exercise is determined by the states of the individual elements (i.e., facts or productions) which make up that exercise. An element can be in one of three states: visited and successfully completed (+), visited and failed (-), or not visited (0). Therefore, each exercise in the lesson can be in one of seven categories as shown in Table 1. The student completes the exercise, and if the student fails, the system completes the appropriate diagnostics. Then the status of all exercises is updated to reflect the current state of the student's knowledge of the lesson. The heuristic search routine selects the next exercise to be presented, based on the information about the status of each exercise in the lesson.

TABLE 1
Possible Exercise Categories Based on the States of the Individual Elements

Exercise Category	Individual Element States
All elements visited and successfully completed	(+)
All elements visited and failed	(-)
All elements visited, and some successfully completed and some failed	(+, -)
All elements not visited	(0)
Some elements not visited, and some visited and failed	(0, -)
Some elements not visited, and some visited and successfully completed	(0, +)
Some elements not visited, some visited and failed, and some visited and successfully completed	(0, +, -)

The exercise selection heuristic is made up of three techniques: (1) handling the apportionment of credit, (2) handling selection procedures in the event of exercise failure, and (3) handling procedures in the event of exercise success.

Apportionment of Credit. Apportionment of credit involves identifying rules which contribute to successful learning and assigning a strength value to those rules. For the case at hand, it is assumed that productions or preconditions that assist in learning are those that have been associated with a successful outcome or a correct response. Therefore, a correct response increases the predictive strength of the rule. Consequently, each time the exercise is successfully completed, its frequency of correct response is incremented by one. Additionally, each time an exercise is visited, its visitation frequency is incremented by one. The inverse of the ratio of frequency of visitation to frequency of correct response defines the predictive strength of the production(s) represented by the exercise. The higher this value, the higher the predictive strength of the productions represented by the exercise. The lower the value, the lower the predictive strength of the exercise. It is also assumed that if a rule made up of composed productions (i.e., where productions are coupled together) is visited, then the experience with this rule also implies experience with the rule's component productions. Consequently, each time such a composed rule is visited, a frequency count should also be added to its component productions. This count is propagated throughout the network. This frequency is called a rule's "coupled frequency". Predictive strength applies to a rule being directly visited in an exercise. Coupled frequency applies to a rule being visited indirectly, and it can measure two things. It is a measure of the associative experience that the trainee has had with rules which are coupled with other rules in the network. It is also a measure of relatedness of a production with other productions. If a rule has a high predictive strength and a high frequency of coupling, then it is a strong rule with a high degree of relatedness to the cognitive simulation model. Predictive strength and coupled frequency can be used by the exercise selection heuristics in determining the selection of an exercise. It is felt that these measures of visitation and coupling assist in implementing the heuristic that one only learns from what one almost already knows and that learning must take advantage of strong prior knowledge.

Exercise Failure. The heuristic search sequence for exercise failure is illustrated in Table 2. The exercise selection heuristic is enabled following two consecutive failures on an exercise. To forgive inadvertent errors in responding, two consecutive failures are allowed. Upon the second consecutive failure, the heuristic searches and selects an exercise comparable to the current active exercise (that exercise which was last failed) which shares: (1) the greatest number of productions or preconditions which were correctly responded to in the current active exercise, and (2) the greatest number of productions or preconditions which were not correctly responded to in the current active exercise. The exercise selected is further constrained because it cannot represent an exercise that has productions or preconditions that are not included in the current active exercise. Additionally, the exercise selected cannot be the current active exercise. The exercise meeting these criteria is the next exercise to be presented.

TABLE 2
Heuristic Search Sequence for Exercise Failure

States of Elements Making Up Exercise				
		Visited		Not Visited
Cycle		Completed (+)	Failed (-)	(0)
1	(+, -)	√	√	
2	(0, +, -)	√	√	√
3	(0, -)		√	√
4	(-)		√	
5	(0)			√

If an exercise which meets these criteria cannot be found, then the second cycle of the search is directed to find an exercise which has: (1) the greatest number of productions or preconditions which were correctly responded to in the current active exercise, (2) the greatest number of productions or preconditions which were not correctly responded to in the current active exercise, and (3) the fewest number of productions not included in the current active exercise. Priority is given to new productions which have the highest predictive strength and the highest frequency of coupling.

If an exercise cannot be found which shares correctly-responded-to productions with the current active exercise, then the third cycle selects an exercise which has: (1) the greatest number of shared productions with the current active exercise which were not correctly responded to, and (2) the lowest number of productions not included in the current active exercise. Again, priority is given to those new productions with the highest predictive strength and highest frequency of coupling. The attempt is always to build on some knowledge which has strength, even if it is indirectly associated with the current exercise.

If an exercise still cannot be found, the fourth cycle selects an exercise which: (1) does not have any shared correctly-responded-to productions with the current active exercise, (2) has the greatest number of shared productions which were not correctly responded to in the current active exercise, and (3) which does not include any new productions not included in the current active exercise. At this point, no other associations can be formed which have any strength.

If an exercise still cannot be found, then the final cycle selects an exercise with the fewest productions not included in the current active exercise, and with the highest predictive strength and the highest frequency of coupling. At this point, the heuristic looks for a new exercise at the lowest level of the AND/OR graph.

To summarize, cycle one and cycle two find exercises which share strong knowledge and weak knowledge with the failed exercise. If these tests fail, then cycles three and four look for exercises which have only weak knowledge in common with the failed exercise. If these tests fail, then one should select an exercise which has indirectly related knowledge with the failed exercise.

Exercise Success. The heuristic search sequence for exercise success is illustrated in Table 3. If an exercise is successfully completed by the trainee, the next exercise on the goal stack is selected. If the goal stack is empty, then one should select an exercise which, compared to the last exercise, shares the greatest number of productions correctly responded to and the fewest number of productions which have not yet been visited and which have the highest coupling frequency.

TABLE 3
Heuristic Search Sequence for Exercise Success

	States of Elements Making Up Exercise		
Cycle		Completed (+)	Not Visited (0)
1	(0, +)	√	√
2	(0)		√

If none can be found (that is, if only exercises with shared productions with the last exercise can be found without any new productions which have not yet been visited), then one should select an exercise with the fewest productions which have not yet been visited and which have the highest coupling frequency. If none exist, then all exercises have been successfully completed. In summary, we have described an intelligent computer aided instruction (ICAI) system which makes empirically valid assumption about knowledge representation. Employing this representation scheme, student diagnosis can provide input to a teaching tactic or teaching strategy which, in turn, can form a modifiable plan of instruction for the individual.

EXPERIMENT

An experiment was designed to evaluate the effectiveness of the cognitive engineering of lesson content and the adaptive sequencing of exercises for learning tactical console procedures. The ICAI system composed of the instructional content and instructional strategy specified above was integrated into the Navy's L-TRAN (Lesson Translator) system.

An example L-TRAN lesson was provided as the test case. The content of the lesson was restructured according to the cognitive engineering process described. The software to implement the adaptive exercise heuristic was developed, and diagnostics were written. The result was three versions of the lesson: (1) the original version currently in use, (2) a cognitive version in which the content was structured according to production rules developed from the cognitive task analysis and was presented in a fixed sequence, and (3) an adaptive version in which cognitively structured exercises were sequenced, based on the individual's performance. The experiment was designed to test the effectiveness of the cognitive and adaptive version as compared to the original version. The experimental hypotheses were: (1) that the cognitive structuring of lesson content would improve performance over the original lesson content, (2) that the adaptive sequencing of the exercises would further improve performance over the cognitive version, and (3) that the adaptive sequencing would be the most efficient of the three versions.

Subjects

Forty-eight enlisted Navy students who had just completed basic training and were ready to begin their school assignment participated in this research on a voluntary basis. There were 40 males and 8 females. The rate of the subjects varied (i.e., 34 Electronics Technicians, 12 Signal Masters, 5 Quarter Masters, 3 Torpedo Men). However, none of the subjects had had any prior experience with tactical training consoles.

Equipment

A Tektronix 4125 graphics terminal and a DEC VT-220 terminal connected to a Micro VAX system, along with a Hewlett-Packard 150II and 915 3A touch-sensitive screen, served as the training workstation. A VCR was also employed to provide necessary background information. The Format-85 authoring language was used to develop the exercises making up the lesson.

The equipment used a software emulation of a Naval Tactical Data System (NTDS) console. The Tektronix 4125 represented the planned position indicator (PPI) and the fixed action buttons (FABs). The VT-220 represented the digital data indictor (DDI), and the HP150C represented the variable action buttons (VABs). The function of the console was to collect and process tactical information, as well as to provide a display of the tactical situation. Functions such as entering, updating, and classifying information on air, surface, and subsurface targets, could be trained. Specifically, the computer-based instruction employed in this experiment involved training the subjects to recognize and use various symbology and console functions associated with tracking a target.

Procedure

Subjects were tested at the Naval Training Systems Center's Human Factors lab. Two subjects were run per day on an individual basis; one was run in the morning and one in the afternoon.

Once the subjects ran through the lesson in its entirety, they were given the performance test. The performance test was scored based on the number of errors made during the test. Errors were automatically recorded by the computer system. In addition, the experimenter observed the subject and manually recorded each performance error made. Subjects repeated this process of running through the lesson and going through the performance test, until the allotted time had expired or until they had attained the performance criterion. Each sequence of lesson and performance test evaluation constituted one trial. The group receiving the adaptive version generally did not repeat the entire lesson on subsequent trials, while the other two groups generally did. Subjects in the adaptive group were routed back to review specific exercises based on their responses to the diagnostics.

Analysis

A single-factor analysis of covariance design was performed on a number of measures to detect any differences in learning that could be attributed to having different versions of the lesson. The covariate was the Armed Services Aptitude Vocational Battery (ASVAB) scores obtained from the Navy. The analysis of covariance was employed to reduce any group differences that may have occurred as a result of varying degrees of aptitude or intelligence. Analyses were performed to determine whether the three groups significantly differed with respect to the number of errors attained on the initial and final performance trials as reflected by the score on the performance test.

Results and Discussion

The goal of the analysis was to determine the effect of lesson structure on procedural learning. The criterion for having learned a procedure is error-free performance of that procedure on the post-lesson simulation test. The criterion for having learned all the procedures presented in the lesson is error-free performance on the entire simulation. Since lesson time was limited to 90 minutes, it was not possible to train each subject to asymptotic performance and then to measure the amount of training time required to reach the error free criterion. The alternative is to look at how close subjects can come to criterion within the allotted time.

In the ET environment, the trainee has a limited time available for training and should therefore get as much as possible out of each session. Efficiency is maximized when the lesson is structured in such a way that each trainee gets the most learning gain in a single pass through the lesson. The three groups were compared for efficiency of training as measured by the mean number of errors made on the first trial of the performance test; that is, after progressing completely through the lesson one time. It must be noted that, for the original and the cognitive groups, the decision to end the lesson was controlled by the subject. They always had the option of going back over any portion of the lesson before ending the lesson and attempting the performance test. However, for the adaptive lesson group, control was determined by the exercise sequencing heuristic; there was no opportunity for the subject to choose to review certain portions of the lesson. An analysis of covariance was performed on the total number of errors made on the first performance test for each subject in each group; the means are shown in Figure 1. A robust effect was found for lesson structure ($p<.0001$).

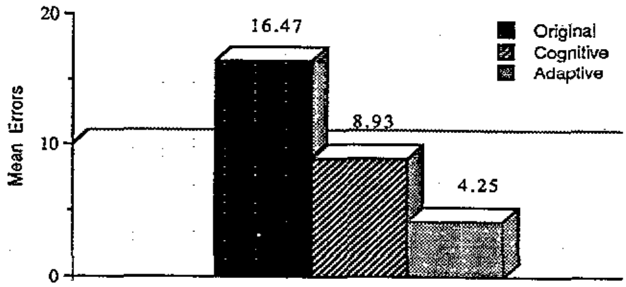

Figure 1. Mean number of errors in each of the three groups on the first trial of the performance test.

After the first trial, subjects were instructed to try to reach the error-free performance criterion within the allotted time period. Figure 2 gives the mean number of errors for each group on the final performance trial of the session. Again there was a significant ($p< .0001$) effect for lesson structure.

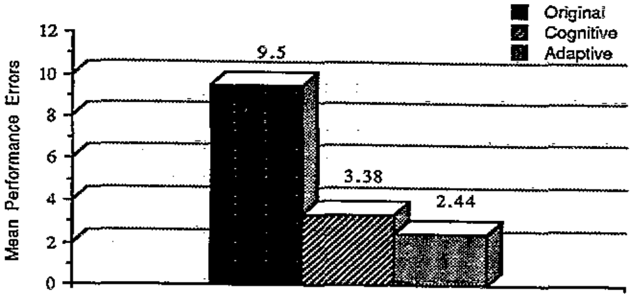

Figure 2. Mean number of errors in each group on the final performance test trial.

Between the first and the last trials, the original and adaptive groups decreased their errors by 42% and 44% respectively, while for the cognitive group, errors were decreased by 62%. With lesson time consistent across groups, the difference in performance was significant and of the same magnitude. For the original and cognitive groups, there is more lesson time remaining after the first trial in which learning can take place than there is for the adaptive group. However, as it turned out, subjects in the adaptive group showed just as much improvement after the initial trial as the original group in less than one third of the time.

CONCLUSIONS

Three general hypotheses were tested: (1) that the cognitively structured version of the lesson would result in improved learning over the original version, (2) that the adaptively sequenced version would result in additional learning gains over the cognitively structured version, and (3) that the adaptive sequencing of the exercises would result in increased efficiency of learning; that is, more improvement in learning per unit of learning time.

The evidence demonstrates that structuring the information to be learned in a computer based lesson according to guidelines developed from a production system model of knowledge representation has a profound impact on learning effectiveness. This impact of lesson structure on procedural learning (that is, the ability to perform procedures correctly) is expected, based on evidence from cognitive learning research. The trainee is presented with an explicitly organized body of knowledge, because the to-be-learned information has been put into a task-goal hierarchy through cognitive analysis, explicit rules have been developed, and exercises have been structured around these rules. And through observing, grouping, and practicing those rules, the trainee develops a procedural model that matches the explicit organization of the instructional material.

When information is not organized in the explicit rule formulation method, learning is slow and difficult. Even after the initial trial, and after exposure to the performance test and the multiple choice test, those subjects with the original lesson version are unable to improve their performance to the level that subjects given the other two versions reach on the first trial.

Procedural learning, as measured by the performance test, is further improved by the adaptive exercise heuristic. It is clear that sequencing takes advantage of the individual's prior knowledge and history of performance on the lesson. By diagnosing the trainee's strengths and weaknesses in terms of specific rules and procedures, the course of the lesson can be built upon the student's partial strong knowledge, helping the student to learn. Additional research is required to further explore the relationship between adaptive sequencing heuristics, the structure of the internal representation, and performance.

Finally, results of this study show that the combination of cognitively structured and adaptively sequenced exercises provides an effective and efficient method of increasing learning in ET systems. The claim of improved efficiency is supported in particular by the finding that subjects with the adaptive version improved performance by 44% after the initial trial, with only about 10 minutes of additional lesson time. The goal of on-board ET is to efficiently and effectively maintain a maximum level of performance and readiness. In the on-board situation, the trainee has already been exposed to initial training and will enter an ET session with varying degrees of development of knowledge structures. The goal of the ET session is to diagnose and correct deficiencies in knowledge and skills as efficiently and effectively as possible. The results of the current study indicate that an adaptive exercise sequencing heuristic is a means for meeting those objectives.

RECOMMENDATIONS

There was a dramatic impact on training gain when the instructional content of the test lesson was restructured using the cognitive engineering process. It is recommended that these results be confirmed and extended through an additional experiment using personnel who possess more background knowledge. Upon confirmation of this effect for cognitive engineering the lesson, it is recommended that: (1) guidelines for authoring L-TRAN lessons using the cognitive engineering process be incorporated into the L-TRAN authoring guidelines and (2) a computer-based authoring system be developed to assist and guide lesson development.

The purpose of the adaptive learning heuristic is to take advantage of each individual's background and training. The current experiment used a homogeneous group of domain-naive subjects. We would expect that further research using subjects with varying degrees of domain knowledge and skills would further highlight the effectiveness and efficiency of adaptive exercise sequencing.

REFERENCES

[1] Anderson, J. R. (1983). The architecture of cognition. Cambridge, Massachusetts: Harvard University Press.

[2] Anderson, J. R. (1986). Knowledge compilation: The general learning mechanism. In R. S. Michalski, J. G. Carbonell, and T. M., Mitchell, (Eds.). Machine learning: An artificial intelligence approach (vol. 2), 289-310. Los Altos, California: Morgan Kaufmann.

[3] Anderson, J. R. Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. Cognitive Science, 13, 467-505.

[4] Kieras, D. E. (1987a). The role of cognitive simulation models in the development of advanced training and testing systems. Office of Naval Research Technical Report No. 23 (TR-87/ONR-23).

[5] Kieras, D. E. (1987b). What model should be taught: Choosing instructional content for complex engineered systems. Office of Naval Research Technical Report No. 24 (TR-87/ONR-24).

[6] Kieras, D. E., & Bovair, S. (1986). The acquisition of procedures from text: A production-system analysis of transfer training. Journal of Memory and Language, 25, 507-524.

[7] Massey, D., Harris, M., Downes-Martin, S., and Kurland, L. (1986) Embedded training technology survey. (Report No. 6062). Cambridge, Massachusetts: BBN Laboratories.

[8] Williams, K. E., and Reynolds, R. E. (1989) Embedded training instructional technology enhancement (NTSC Tech. Rep. 89-034). Orlando, Florida: Naval Training Systems.

[9] Williams, K. E., Reynolds, R. E., Carolan, T. F., Anglin, P. D., and Shrestha, L. B. (1989) An evaluation of a methodology for cognitively structuring and adaptively sequencing exercise content for embedded training. (NTSC Tech. Rep. 89-035), Orlando, FL: Naval Training Systems Center.

ABOUT THE AUTHOR

Dr. Kent E. Williams received his Ph.D. in Cognitive Psychology from the University of Connecticut in 1976. Since that time, Dr. Williams has been conducting applied and basic research in cognitive science as it relates to training and simulation. Dr. Williams is currently the Research Manager for the Institute for Simulation and Training at the University of Central Florida, Orlando, Florida. Dr. Williams' business phone number is (407) 658-5025.