U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1864

The Training, Retention, and Assessment of Digital Skills: A Review and Integration of the Literature

Gregory A. Goodwin
U.S. Army Research Institute

November 2006

U.S. Army Research Institute for the Behavioral and Social Sciences

A Directorate of the Department of the Army Deputy Chief of Staff, G1

Authorized and approved for distribution:

STANLEY M. HALPIN
Acting Technical Director

MICHELLE SAMS
Acting Director

Technical review by

Brooke B. Schaab, U.S. Army Research Institute
William R. Sanders, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Research Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-MS, 2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926

FINAL DISPOSITION: This Research Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Research Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
November 2006	Final	August 2005 - October 2006

4. TITLE AND SUBTITLE The Training, Retention, and Assessment of Digital Skills: A Review and Integration of the Literature	5a. CONTRACT OR GRANT NUMBER
	5b. PROGRAM ELEMENT NUMBER 622785
6. AUTHOR(S) Gregory A. Goodwin (U.S. Army Research Institute)	5c. PROJECT NUMBER A790
	5d. TASK NUMBER 215
	5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Infantry Forces Research Unit P.O. Box 52086 Fort Benning, GA 31995-2086	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Research Report 1864

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Subject Matter POC: Gregory A. Goodwin

14. ABSTRACT *(Maximum 200 words)*:

Skill training, retention, and assessment research from academia, industry, and the government is reviewed and its implications for digital skill training are discussed. Factors that affect digital skill retention are procedural (e.g., the timing and sequencing of training trials, training techniques employed), individual (e.g., intelligence or background knowledge), and task related (e.g., the number of steps per task, the cognitive difficulty of selected steps). Some of these factors have also been shown to impact digital operator training, but additional research needs to be done. Few experiments have made broad comparisons across all variables or have effectively quantified ways to evaluate or design effective digital training. Recommendations about how to modify specific digital courses should be validated. Much less work has been done in the area of collective digital training despite the fact that this comprises the majority of the training prescribed in the Army Digital Training Strategy. Research on the assessment of digital skills has produced a number of tools and procedures for assessing digital skills for both individual and collective training. Many of these tools have been well received by units who have used them but the tools still need to be formally validated.

15. SUBJECT TERMS

operator skill, employment skill, skill decay, skill acquisition, digital skills, Army Battle Command System, assessment

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT	17. ABSTRACT	18. THIS PAGE			Ellen Kinzer
Unclassified	Unclassified	Unclassified	Unlimited		Technical Publication Specialist (703) 602-8047

Research Report 1864

The Training, Retention, and Assessment of Digital Skills: A Review and Integration of the Literature

Gregory A. Goodwin
U. S. Army Research Institute

Infantry Forces Research Unit
Scott E. Graham, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926

November 2006

Army Project Number 622785A70	**Personnel Performance and Training Technology**

ACKNOWLEDGMENTS

The author would like to express his gratitude to Dr. Bruce Leibrecht, Mrs. Karen Lockaby, and Mr. Christopher Strauss of Northrop Grumman for their help in gathering some of the reference materials for this review.

THE TRAINING, RETENTION, AND ASSESSMENT OF DIGITAL SKILLS: A REVIEW AND INTEGRATION OF THE LITERATURE

EXECUTIVE SUMMARY

Research Requirement:

The Army has adopted a suite of digital command and control systems collectively known as the Army Battle Command System (ABCS). Although the systems that make up the ABCS allow leaders to send orders, maps, and other information across the battlefield, they have also created new challenges to warfighters. Chief among these challenges is training personnel to use them to their fullest advantage.

Users of these systems consistently report that individual operator skills are perishable and require frequent use to maintain. Furthermore, training to employ these systems can be a challenge. Units often lack the necessary training and evaluation guides and often do not understand how to plan effective training events.

Scientists at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) have been doing research to address these ABCS training, retention, and assessment issues. Although parts of this work have been summarized in reviews published over the last decade, there has been no comprehensive review of this body of research. Furthermore, research from academia and other agencies has begun to focus on digital training and retention. The present report is an effort to both review research on the training, retention, and assessment of digital training and to identify directions for future work in this area.

Procedure:

Citations were derived from both governmental and academic literature databases. Online search engines and subject matter experts were employed to help identify relevant articles. Research on the retention of digital skills was integrated into the larger body of work on skill retention to help organize the former and identify gaps in the research on digital skills. The research on digital training and the assessment of digital skills was analyzed from the standpoint of providing recommendations for both training and future research.

Findings:

Factors found to affect individual skill acquisition and decay include procedural variables (e.g., number of training trials, retention interval, and training approaches), task variables (e.g., task complexity or the number of steps in a task), and individual variables (e.g., intelligence and background knowledge).

The review of this literature found several practices to improve digital skill training and retention (e.g., breaking down material into reasonable chunks, employment of various training

principles such as context interference or self-guided training) as well as practices to avoid (e.g., overtraining during initial training or spacing of training trials). Assessment tools and techniques are available for researchers and operational units, but many of these tools still need to be validated. Future work should leverage this body of knowledge and experience to improve the training, retention, and assessment of digital skills.

Utilization and Dissemination of Findings:

This report was initiated in response to a request from the Training and Doctrine Command Capabilities Manager (TCM) Stryker-Bradley office at Fort Benning. The request was for guidance on how to reduce digital skill decay and better prepare Stryker Brigade Combat Teams (SBCTs) to use their digital ABCS systems. The findings of this literature review were briefed to the TCM Stryker-Bradley office along with other findings from a work package on improving digital skill training and retention.

THE TRAINING, RETENTION, AND ASSESSMENT OF DIGITAL SKILLS: A REVIEW AND INTEGRATION OF THE LITERATURE

CONTENTS

LIST OF FIGURES

Introduction

Digital systems are increasingly used to communicate and process battlefield intelligence for the Army. Systems have been developed for fire support (Army Field Artillery Automated Tactical Data System – AFATDS), operations (Maneuver Control System – MCS), intelligence (All Source Analysis System – ASAS), and vehicle tracking and communication (Force 21 Battle Command Brigade and Below – FBCB2), among others. Collectively these are known as the Army Battle Command System (ABCS), and they provide leaders with the ability to send orders, reports, graphics, and other data across a wireless digital network that covers the battlefield. Although this technology has the capability of greatly facilitating information flow and analysis, it has introduced new challenges to the warfighters who use it.

Chief among these challenges is training and sustaining proficiency on ABCS systems. This challenge has been a surprise to some who thought that Soldiers would be able to train themselves to use these digital systems. Such an expectation is undoubtedly based in the widespread preference among both civilians and service members to learn new software programs, including ABCS systems, by "playing with the software" (Dryburgh, 2002; Schaab & Dressel, 2003). Research done within ARI and elsewhere, however, has shown that this is an inefficient and ineffective means of digital training (Charney, Reder, & Kusbit, 1990; Dyer, Singh, & Clark, 2005).

Another training challenge stems from the pace at which software for digital systems is constantly being revised and patched. These updates make the software functionality a moving target for training developers. Functions, buttons, and menus are always being added, subtracted, and moved, meaning that there is a strong likelihood that the system a Soldier trains on may not be the system that he or she uses in the unit.

Once trained, Soldiers and leaders have found that these so called *digital skills* are highly perishable. Much of the evidence for this comes from anecdotal reports by various leaders (e.g., Lynch, 2001) or from analysis of training exercises like the Advanced Warfighter experiment ("Advanced warfighter experiment focused dispatch final report", 1996). This belief was also conveyed by digital leaders who participated in ARI's *Managing at the Speed of Change in Force XXI* project (Johnston, Leibrecht, Holder, Coffey, & Quinkert, 2003). Surprisingly there has been little empirical data to back this up, and one ARI report questions the validity of the assumption (Schaab & Moses, 2001).

The likelihood that these skills are perishable agrees with what psychologists have known about discrete procedural skills (a category into which most digital skills fall) since the 1950's (Adams, 1987), namely that they are rapidly forgotten. For example, research on aircraft pilots has shown that their recall of discrete cockpit procedures can decay to unsafe levels within a matter of weeks in the absence of continued training (Schendel, Shields, & Katz, 1978).

In addition to sustaining digital skills, other training challenges have surfaced including training on software upgrades, incompatibility between different systems, the availability of digital training facilities, and the challenges of learning to process and manage large volumes of information (Johnston et al., 2003; Schaab & Dressel, 2003). These challenges have attenuated some of the expected benefits of digital command and control systems on the battlefield (Lynch, 2001).

To help the Army overcome some of these problems, ARI has conducted research on digital skill training, including work on individual operator training (e.g., Sanders, 1999; Schaab & Dressel, 2001), proficiency measurement (e.g., Leibrecht, Lockaby, Perrault, Strauss, & Meliza, 2006), and training on future digital systems (e.g., Dyer et al., 2005; Lickteig, Sanders, Lussier, & Durlach, 2004). Although there have been reports summarizing portions of this research (Schaab, Dressel, & Moses, 2004; Throne & Lickteig, 1997), there have been no comprehensive reviews of this body of work. The present report is an effort to integrate research on digital skill training, retention, and assessment from academia, government, and industry in order to develop recommendations for Army digital trainers and identify directions for future research.

Definition of Skill

Before discussing the research on digital skills, it is worth taking the time to define the concept of skill. Pear (1927), a British psychologist, provided one of the earliest definitions of the term in an American journal of industrial psychology. He felt that skills had several features that distinguished them from aptitudes, or habits. First, he said that a skill must be learned. Thus, walking on a tightrope would be a skill, whereas, walking on the ground would not. Additionally, he said that skills required an integration of many parts into a component whole. Thus, juggling several balls would be a skill, whereas tossing a single ball in the air would not. Finally, he said that skills were primarily motor behaviors.

The problem with this last requirement is that any skilled action relies on both cognitive and motor output, although the balance often leans towards one or the other (Adams, 1987). For example, a mathematician must rely on motor output to write the solution to a problem, but this is only a trivial part of his or her skill. A tennis player or a marksman, on the other hand, must make mental calculations, but their skill depends much more heavily on practiced motor output. A compromise is to recognize both cognitive and motor skills. Although at least one researcher (Adams, 1987) has somewhat recently insisted that the term skill be reserved for predominantly motor tasks, the phrase *cognitive skill* is widely used throughout the psychological literature (a recent search on PsychINFO and PsychARTICLES for the phrase turned up 2577 references).

This distinction between cognitive and motor skills is not trivial. For example, it is conceivable that the size of an individual's short term memory register might influence performance on predominantly cognitive skills yet have little effect on predominantly motor skills. Other differences may exist as well. For example, Schendel et al. (1978) concluded that distributed practice has no impact on the retention of motor skills, whereas Fendrich (1988) concluded that distributed practice has a major impact on the retention of cognitive skills.

By allowing both cognitive and motor acts to be considered skills, a wide range of human activity can potentially be considered a skill, and one has to wonder whether this definition of skill is too inclusive to be useful. Of course, there is still the criterion that a skill be a complex, learned behavior, but even this restriction may not be all that useful. The problem is that there is no objective means of dividing acts that are sufficiently complex to be considered skills from those that are not. In fact, even in the literature on cognitive skills, simple verbal learning tasks, such as memorizing word lists, are often studied as exemplars of cognitive skills (e.g., Fendrich et al., 1988; Healy, Meiskey, Fendrich, Crutcher, & Little, 1988). Rather than abandon the term, skills are thought of as existing along both a motor-cognitive continuum and a simple-complex continuum for the purpose of this review. Distinctions along these two continua are discussed whenever relevant.

As this is a review of *digital* skills, it is necessary to discuss their place along these continua. Digital skills are those needed to use software running on a computer. At the current time, this involves some combination of data entry and the execution of commands through a graphical user interface (GUI). From the standpoint of the Army, digital skills are increasingly used on the battlefield by users of the ABCS.

Digital skills are discrete, multi-step procedures (i.e., navigation through a series of menus and submenus to set parameters and execute commands). Although there is a motor component to digital skills, namely moving a pointing device or operating a touch-screen, the motor skill level required is fairly minimal (Sanders, 1999). Thus on the continuum of motor to cognitive skills, these skills are more cognitive than motor.

The place of digital skills along the simple-complex continuum depends on the specific skill. Digital skills can be either individual operator skills or collective employment skills. Individual operator skill is the ability to get a given digital system to perform one of its functions. Collective employment skill is the ability of a group of Soldiers or leaders to determine when and how to use the functions of a network of ABCS systems during the course of an operation/exercise. It can be seen that they range from relatively simple individual skills (clicking a single button) to very complex collective skills (employing multiple digital systems to command and control a brigade on the battlefield). Thus, digital skills span the simple to complex continuum.

Overview

The following review is divided into three sections covering research on digital skill training, retention and assessment. In the first section, covering digital skill training, research is described that examines the ways in which theories of learning have been applied to improve digital skill acquisition. In addition, research on improving training on information management is summarized.

In the second section of this review, research on digital skill retention is discussed in the larger context of skill retention research. Using this context makes it possible to identify areas where research on digital skills is lacking. In areas where research on digital and non-digital skills parallel one another, comparing the findings makes it possible to determine how readily findings can be generalized across these two categories.

Research on training and retention are, to some degree, difficult to separate in that retention is partially a function of training. The distinction drawn in this review is that research concerned with the content of training is discussed in the section on training whereas research concerned with the parameters of training (i.e., the duration, frequency, retention interval, etc.) is discussed in the section on retention. Another distinction between these two areas is that training researchers typically focus on improving the rate of acquisition while retention researchers are more concerned with the durability of the trained skill. These two objectives are sometimes mutually exclusive as will be discussed below.

The final section of this review covers research on digital skill assessment. Assessment is important for a couple of reasons. First, the measurement of digital skill proficiency underlies any research on skill acquisition and retention. Without reliable measures, it would be impossible to determine the effectiveness of any experimental manipulation. Second, the measurement of digital skills has a very practical use. It is important to unit leaders who need to know the frequency and type of training to schedule for their unit. This section reviews research on skill assessment of individuals or groups using existing ABCS systems as well as a notional Future Combat System.

Digital Skill Training

Training principles derived from theories of learning.

Much of the research done on training principles can be understood in the context of psychological theories of learning. The three most relevant theories are behaviorism, cognitivism, and constructivism. Behaviorists view the learning process as one in which actions or cognitions are reinforced or punished and thereby persist or extinguish. Learning is therefore a passive process. Cognitive psychologists on the other hand, see the learner as an active information processor. Constructivism (a derivation of cognitivism) sees the learner as one who learns by constructing concepts based on experience (Sanders, 2001). These different theories of learning suggest a variety of training principles that can influence the makeup of what has so far been referred to as a training trial.

Sanders (2001) talks about how to apply these theories to the training of digital systems by using FBCB2 as a case-in-point. For example, training derived from behavioral principles includes doing an FBCB2 task analysis and modeling digital tasks for students. Training derived from cognitive principles would include creating outlines, advance organizers and summaries of the material as well as using mental visualization or mnemonics to facilitate encoding. In fact, cognitive training principles were used to develop computer based instruction modules for some FBCB2 operator tasks (Deatz & Campbell, 2001). Finally, training derived from constructivist principles would include incorporating realistic scenarios in training and the use of instructors as coaches to guide students as they develop their own solutions to problems.

An example of the application of cognitive psychology principles to digital training development is found in Dyer and Salter (2001) who examined the effect of varying the demand placed on working memory during training. Working memory demand was varied by creating lessons in which large (from 9 to 18 chunks) or small (from 3 to 9 chunks, lessons were paired so that the small always had fewer chunks than the large) amounts of information were presented before participants could apply them. The results favored training with smaller, more manageable chunks of information. This finding suggests that once working memory is at capacity with new information, individuals must have an opportunity to consolidate that information for it to be retained.

Another recent example of the application of cognitive skills to the learning of computer skills is an experiment by Davis and Yi (2004). Participants were trained to perform several procedures in Microsoft Excel using a method of mental imaging called symbolic mental rehearsal (SMR). To perform SMR, participants break the steps in a task down into subgroups and assign labels to them (i.e., copy the formula into all rows). Participants next wrote the labels down and practiced mentally rehearsing the sequence of labels. SMR performance was measured immediately and 10 days after initial training. At both times, SMR was better than a control (watching a demo and then practicing on their own). This was not related to affective response (i.e., SMR wasn't more fun or motivating than the control condition). The authors carefully controlled for effects of different instructors, gender, age, and computer and spreadsheet experience both by showing no main effects and by using those variables as covariates in the analysis.

An example of the application of constructivist principles to train digital skills is provided in a report by Schaab and Dressel (2001) in which Military Intelligence officers were trained to use ASAS with either traditional (i.e., guided demonstration) or constructivist techniques. In the constructivist condition, participants were given minimal instruction and then worked in groups on a series of practical exercises (PEs). In the traditional training group, students followed an instructor who provided a lecture, did a demonstration, and then gave them a PE to work through. Both groups had equal training time. Although both groups performed equally well on the standard final exam, the guided exploration group performed better on a novel PE and reported lower levels of cognitive load on all tests.

The interaction between training and aptitude. In the 1960s and 1970s the research described in this section would have been subsumed under the larger battle among theories of learning (i.e., behaviorism or cognitivism) but in the current zeitgeist, there is little discussion of which theory of learning is superior. Rather, researchers try to identify the conditions under which various techniques work best. Along these lines, Clark and Wittrock (2000) conceptualized common approaches to training, whether digital or non-digital, along a single continuum of instruction ranging from external (instruction is driven by environmental factors) to internal (instruction is driven by factors that are internal to the learner). The four types of training that they place along this continuum are: receptive (teaching by telling), behavioral (teaching by demonstration, practice, and feedback), guided discovery (teaching by problem solving), and exploratory (teaching by exploration). Clark and Wittrock emphasize that the important question is not which approach is best but which approach is best for certain individuals. For example, they believe the receptive approach works well for novices but the exploratory approach works better for individuals with critical background knowledge and the motivation and metacognitive skills to train themselves.

Mayer, Greeno, and their colleagues have done much work examining the interaction between ability and training approach (reviewed in Fendrich et al., 1988; Mayer, 1975) though they have not specifically looked at training on digital systems. In one experiment Greeno looked at the performance of participants in different ability groups when receiving rule based training (equivalent of Clark and Wittrock's behavioral approach) or discovery training (equivalent of Clark and Wittrock's guided discovery approach). Participants learned to solve probability equations. Prior to the experiment, all participants were given a pre-test of their understanding of probabilistic concepts and equations. In the discovery condition, participants were given an equation and then asked to generalize the equation to solve other problems with little instruction. In the rule group, participants were given an equation and then explicit instructions on how to solve the additional problems. Participants were all tested after training. Pre-test scores mattered for individuals in the discovery condition where those who scored high on the pre-test did better on the final test than those who scored low. Pre-test score made no difference for individuals in the rule condition.

A report by Baldwin et al. (1976) shows a similar interaction between ability and training method examining air defense artillerymen. In this experiment, participants were given aircraft recognition training. They were divided into low, intermediate and high ability groups based on their general technical score (a combination of the verbal and math scales of the ASVAB). Pre-test and post-test (taken one week after training was completed) measures of aircraft recognition were taken and participants were divided among self-paced (worked at own rate with a buddy using flashcards) and group-paced (observed slides in a classroom with instructor asking questions) training. Under group-paced training, the low ability group showed greater gains than the intermediate ability group but under self-paced training the opposite was true. The high ability group displayed high pre-test performance (87% accuracy) and did not significantly improve under either condition.

An investigation of digital training by Dyer et al. (2005) is consistent with the work of Mayer, Greeno, and Baldwin. In this experiment, participants in basic training (OSUT) and the Infantry Officer Basic Course (IOBC) were given training on using a digital map display. Participants were assigned to one of four training conditions plus a condition where they could choose their own method of instruction. One condition was a free exploration condition. Three conditions were types of guided exploration. In the first, participants performed exercises with feedback. In the second condition, they received formal training and then were able to explore the system on their own. The third guided exploration condition was a combination of the other two. Results showed that OSUT participants (new enlisted Soldiers typically without a college degree) scored worse in the free exploration condition than in any of the other conditions; however, the IOBC participants (new officers with a college degree) scored equally well regardless of condition. Considering that IOBC participants had a longer and more advanced formal education than OSUT participants, it could be argued that the former group had a higher ability, by virtue of their experience, to learn in the training environment of this experiment. If this interpretation is valid, this experiment suggests that, as with other types of training, digital training should be tailored to the ability level of the students.

Training principles for digital training. A number of other reports of digital training have compared the effectiveness of the training approaches outlined by Clark and Wittrock (2000), although the reports do not always consider the modulating effects of ability. Much of this research is discussed in the review and annotated bibliography of training computer skills by Throne and Lickteig (1997) and some have been published since (e.g., Davis & Yi, 2004; Dyer & Salter, 2001; Dyer et al., 2005; Schaab & Dressel, 2001). Making comparisons across all of these reports is somewhat difficult because there is no real standardization of training approaches tested. Nevertheless, some general conclusions can be drawn from this body of work. It is important to keep in mind that most of these studies have used initial acquisition and not long term retention to determine training effectiveness.

The first conclusion is that completely unguided exploration in which participants are allowed to explore the software but are given no training materials or exercises other than a users manual is the least effective means of training digital skills (Charney et al., 1990; Czaja, Hammond, Blascovich, & Swede, 1986; Dyer & Salter, 2001; Dyer et al., 2005). This approach does have one advantage in that some reports say that it takes the least amount of time (Charney et al., 1990; Dyer & Salter, 2001).

Another conclusion is that computerized tutorials are not as good as behaviorally modeled procedures (Czaja et al., 1986; Gist, Rosen, & Schwoerer, 1988; Gist, Schwoerer, & Rosen, 1989). This is true whether the behavior is modeled by a live demonstrator or a videotaped demonstrator. One report indicates that a computerized tutorial is better than text only training during acquisition but not recall (Palmiter & Elkerton, 1993). In this experiment participants were trained to do a variety of procedures ranging from 3 to 12 steps in a Macintosh program called Hypercard. Participants who were trained using a computer animated demonstration performed better than participants who trained with a text only description immediately after training but one week later, those who were in the text only condition showed better recall. Once again, this highlights the importance of looking beyond an immediate test of performance when evaluating the relative strengths of different training approaches. One caveat of this general conclusion about the inferiority of computerized tutorials is that in the decade and a half since most of this work was done, computerized tutorials have changed a great deal. Videos that provide behavior modeling are sometimes included as part of the tutorial, and most computerized tutorials today include practical exercises. It is therefore not clear that this conclusion would stand up today.

Finally some variation of guided exploration (a constructivist technique in which students are given minimal instruction and then must work through a series of exercises) is better than other conditions tested such as unguided exploration, behavioral modeling, computerized tutorial, or classroom instruction alone (Carroll, Mack, Lewis, Grischkowsky, & Robertson, 1985; Charney et al., 1990; Frese, Brodbeck, Heinbokel, Mooser, & et al., 1991; Schaab & Dressel, 2001). The experiment by Frese et al. made an interesting comparison between behavioral modeling in which participants were given incomplete information on the steps to be followed and were encouraged to figure out errors on their own (error training group) and behavior modeling in which participants were told exactly what to do and were immediately corrected when they made mistakes with no further explanation (error avoidant group). The error training group could recall more of the steps to the procedures and was better at spotting mistakes than the error avoidant group.

One exception to these findings was reported by Simon and Werner (1996). In their research, they found that guided exploration was not as good as behavioral modeling in conjunction with guided exploration. These authors compared three modes of training on an automated data processing system used by the Naval Construction Force. The first mode was classroom instruction in which participants observed a lecture and slideshow demonstrating procedures but did not have an opportunity to perform the tasks prior to the test. The second was a guided exploration condition in which participants were seated at the computer and had a series of exercises they had to work through. The third condition was behavior modeling plus guided demonstration in which participants were trained as in the classroom instruction condition but had the opportunity to model the behaviors and ask questions of the instructors. In addition, participants in this last condition were given the materials of the guided exploration group and were encouraged to work through the exercises.

Participants were given both a knowledge test and a performance test both immediately and one month after initial training. The behavioral modeling plus guided exploration group performed better than all other groups on both tests at both time points. In addition, the guided exploration group did better on the performance test than the classroom instruction group at both time points and did better on the knowledge test at the one month follow up. Interestingly, the classroom instruction group did better on the knowledge test immediately after training.

In summary, a variety of training approaches show promise for improving standard operator level training on digital Army systems. For example, training instructors to evaluate the way course content is subdivided will prevent students from being overloaded with information. Use of SMR may also help students better encode information. Finally the use of guided exploration in conjunction with behavioral modeling and the incorporation of constructivist training techniques is likely to improve the acquisition of digital skills.

Research questions still remain regarding the impact of different theoretical approaches on digital skill retention. Specifically, the benefit of different training approaches over time is still a relative unknown. Few of the reports reviewed in this section examine the retention of digital skills for longer than one month and most rely on data from a test taken immediately following training, so from the standpoint of long-term retention, it is difficult to know which approach is most beneficial.

Another problem is that the differences between these approaches are more qualitative than quantitative and so when it comes to specific questions of course design, it is difficult to make recommendations. For example, as described above, unguided exploration is generally worse than guided exploration but how much guidance is needed to show a benefit? Similarly, what is the optimal balance of behavioral modeling and guided exploration? Because none of these variables are easily scaled, it is hard to make a priori recommendations on how to fine tune any specific program of instruction.

Given these gaps in our knowledge about the impact of training approaches on long-term retention and about the specific ways to fine-tune course POIs for optimal effect, the training benefit of any recommended changes should be empirically validated before they are adopted on a wide scale. Excellent examples of such an empirical validation are the research described above by Schaab and Dressel (2001) or Dyer and Salter (2001). By doing this sort of analysis, measurable changes in training effectiveness can be weighed against the cost of implementing the changes in the POI. This approach to modifying digital system POIs will insure that updates are made in the most cost-effective manner.

Information Management

Thus far in the discussion of system training, the research reviewed has focused exclusively on the operation of digital systems. In order to employ these systems effectively, however, individuals must be trained on more than just *buttonology,* they must be trained to employ their systems. One of the most critical employment skills leaders with digital C2 systems need to posses is how to manage the volumes of information available to them.

The potential challenges of increases in information flow brought about by digital C2 systems have been recognized in a number of reports (Archer, Warwick, McDermott, & Katz, 2003; Leyden, 2002; Moses, 2001). These reports warn of the possibility that too much information may reduce rather than increase situational awareness. In fact, an early Advanced Warfighter Experiment demonstrated that a digitally equipped task force did not make decisions in less time than non-digital counterparts. This was largely because the intelligence units had a difficult time inputting and analyzing the volumes of information they received (Swinford, 1997).

In an experimental examination of this problem, platoon leaders (PLs) were trained to use a simulated digital command and control system (Lickteig & Emery, 1994). The PLs received various types of messages and had to decide what to relay to higher and lower echelons. The experimenters varied the relevance of the information and the volume of the information given to the PLs. In the high relevance condition, all messages were relevant to the PL's sector. In the low relevance condition, only one third of the messages pertained to the PL's sector. In the high volume condition, PLs received messages every 26 seconds and in the low volume condition, every 60 seconds.

High volumes of information led PLs to be both less aware of enemy locations and less likely to send relevant messages to their superiors. In the low volume condition, PLs relayed significantly more irrelevant messages. In the low relevance condition, PLs took more time to read and display messages and were more likely to relay irrelevant messages to superiors and subordinates. In addition, PLs in this condition were less accurate in their understanding of enemy and friendly strength and status. Contrary to expectations, PLs in the low relevance condition were just as accurate in their knowledge of enemy and friendly locations on the battlefield (Lickteig & Emery, 1994).

These reports clearly indicate that the large volumes of battlefield information have the potential to adversely affect leaders' decision making abilities, but little research has been done to determine how to best train leaders to manage large volumes of information. One report, currently in preparation, examines decision-making in an information-dense environment and also examines training to overcome decision-making errors (Folds, Blunt, & Stanley, in preparation).

In this series of four experiments, individuals and teams of Reserve Officer Training Corps cadets and college undergraduates received unfiltered information (e.g., a 9-1-1 call with someone reporting gunfire heard, live footage from a news outlet showing a burning vehicle, etc.) through a software interface, and they had to piece these data together to determine whether a critical incident had occurred. There were six types of critical incidents participants were instructed to look for including sniper fire, armed mobs, and credible evidence of terrorist activity.

The authors of this report were interested in the effect of information volume (the number of events presented during a trial) and density (the ratio of relevant to filler events during a trial) on the ability of the participants to identify and report critical incidents. The experiments were designed so that the frequency of certain types of decision-making biases could be recorded. In two of the experiments, anti-bias training was administered to determine whether training could reduce the occurrence of errors. The biases examined are described below:

- *Vividness* – Information derived from subjective interpretation of pictures or sounds may be more influential than information available in other formats.
- *Absence of evidence* – The fact that evidence is missing, when logically it should be present, is not properly considered.
- *Availability* – Decisions are often influenced by recent events or well-known conjectures that provide convenient explanations for observations.
- *Over sensitivity to consistency* – Multiple reports that in fact are derived from a single source may be treated as though they are independent confirmations of the observation.
- *Persistence of discredited information* – Information that was deemed relevant often persists even after it has later been discredited.
- *Randomness* – In general there is a bias against defining something as random. Often people will impose a causal relationship where none really exists.
- *Small sample* – Evidence from small sample sizes is given equal weighting to evidence from larger sample sizes.

Once participants had sufficient supporting data for a critical incident, they would include the relevant data files in a report. Reports could either be of critical incidents that should be reported (i.e., the data supported an incident of terrorist activity) or filler incidents that should not be reported (i.e., the data did not clearly support an incident of terrorist activity). Some filler incidents were constructed to look like terrorist activity but they lacked critical indicators to warrant reporting them (e.g. a vehicle is reported burning on the roadside but no other evidence exists that it was related to an attack). These types of incidents were called false alarm opportunities (FAOs). The information about these FAO incidents was designed with the biases in mind so that the reporting of the incidents served as an indicator that the participant had fallen victim to a particular bias.

Across all experiments, participants were significantly more likely to report critical incidents than FAO incidents suggesting that they could discriminate between the two. Information volume and density did not have an effect on the reporting of critical incidents but sometimes affected the reporting of FAOs. The authors concluded that their manipulation of these variables was not robust enough to produce a substantial effect.

Although no single bias was the most prevalent in every experiment, the oversensitivity and vividness biases were the most frequently occurring across the series. When teams were tested, they seemed to be more susceptible to the oversensitivity bias. This may have been because the team members were not as likely to communicate the sources of their information to each other. Thus, what would have appeared to the leader to be multiple confirmations of an incident, in fact, were repetitions of the same piece of information.

Initially the anti-bias training was not effective. In that experiment, participants with training were significantly better at identifying information traps than the untrained participants, but the training did not improve the reporting of either critical events or FAOs. After the training was refined in a subsequent experiment, participants avoided reporting almost all FAOs, leading the authors to conclude that the training was highly effective.

The authors describe several critical lessons in designing anti-bias training. First, they recommend multiple experienced reviewers read and critique the training to make sure it is logical and clear. Second, they suggest the training be as close to the real-world task as possible. Third, they suggest that the content of the training be free of any jargon or abstraction and provide multiple clear examples and opportunities for practice. Finally, they suggest that the training be thoroughly tested on individuals not involved in the training development.

Research in the area of information management has uncovered the types of errors and biases that can impact decision-making in an information-dense environment, and it has shown that training has the potential to mitigate some of these problems. What is not currently known is the extent of these problems in existing Army units that employ digital C2 systems. Many leaders in these units have extensive combat experience in units equipped with digital systems, but the extent to which they have encountered and/or overcome these problems in this real-world setting is unknown.

Digital Skill Retention

Factors That Affect Skill Retention

As Hagman and Rose (1983) point out in their review of the retention of military skills, there are three ways to improve retention: improve training, modify the task, and select persons with certain abilities or aptitudes. Research on skill retention has focused on three corresponding categories of variables. One category has to do with the properties of the training and testing (procedural variables). Examples of these variables include massed vs. distributed training,

training to proficiency or to mastery, and duration of the retention interval. Another category has to do with properties of the task (task variables). Examples of these variables include, number of steps involved in the task, the complexity of the steps, and whether the task is continuous or discreet. The final category is related to the characteristics of the individual being trained (individual variables). Examples include the aptitude of the individual and whether or not the individual has certain background knowledge or expertise.

There are differences in the way the effects of these variables have been interpreted and this has sometimes led to confusion in the literature. The confusion arises because some authors feel that a given variable affects retention by changing the score on the final skill retention measure, whereas others feel that a given variable affects retention by changing the decay rate. Statistically this is a disagreement about whether a main effect of the factor or an interaction effect between the factor and time constitutes an effect on skill retention.

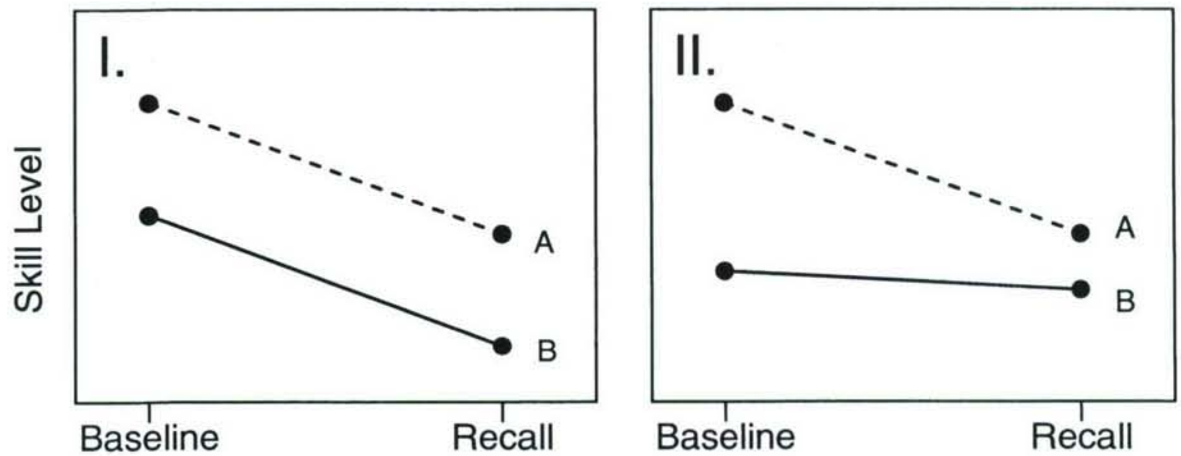

Figure 1. Hypothetical retention curves showing effects of a variable on skill decay. In panel I, both groups forget the same amount of material but group A performs better than group B on the recall test. In panel II, group B forgets less than group A but scores worse on the recall test.

Figure 1 illustrates this problem. In both panels, skill level is measured following training (baseline) and following some retention interval (recall). The participants are divided into two groups (A and B) that are administered one of two levels of some hypothetical variable. In panel I, group A performs better at both measurement times. Researchers concerned with absolute score at recall would interpret this to mean that the factor significantly affected retention because the scores on the recall test were higher in group A than group B. Researchers concerned with decay rate would come to the opposite conclusion and say that because both groups forgot equal amounts of material, initial learning rather than retention of the material was affected by the factor.

In panel II these two groups of researchers would agree that retention was affected but disagree as to which group showed better retention. Researchers concerned with absolute recall would say that the factor improved retention in group A because it had a better score at recall. Researchers concerned with decay rate, on the other hand, would say retention was improved in group B, because that group forgot significantly less than group A. To avoid this confusion throughout the review, a distinction will be made between performance at recall and decay rate.

Procedural Variables

Retention interval. It perhaps goes without saying that the longer an individual goes without practice, the more forgetting will take place. The slope of the forgetting curve for verbal tasks is steepest initially and then its slope declines with time (Ebbinghaus, Ruger, & Bussenius, 1913). This type of forgetting is assumed to apply to motor skills as well (Schendel et al., 1978) but little research has been done to verify this. In their meta-analysis of skill decay (which included both motor and verbal skills), Arthur, Bennett, Stanush, and McNelly (1998) find that retention interval and recall are negatively correlated ($r = -.51, p < .05$). In that analysis, retention interval was broken down into 8 intervals ranging from one day to greater than one year. Although they find a relatively strong correlation between time and skill decay, the authors used a linear correlation statistic and so they may have underestimated the relationship between these variables.

A greater retention interval is not always associated with greater skill decay. In an investigation of the reacquisition of combat engineer procedural skills during an Individual Ready Reserve (IRR) train-up, the period of separation from active duty did not affect skill decay (Wisher, Kern, Sabol, & Farr, 1994). Participant data was placed into one of three groups: those separated for 24 months or less, those separated for 25 to 48 months, and those separated for more than 48 months. The most rational explanation for this apparent lack of an effect of retention interval is that the forgetting curve had flattened out by 24 months and the analysis was not powerful enough to detect small changes in skill decay.

Retention Interval and digital skills. Despite a common perception that digital skills are highly perishable (Johnston et al., 2003) there is very little empirical data documenting the rate or extent of skill decay. Schaab and Moses (2001) did report non-experimental results in a cross-sectional sample suggestive of some skill decay among ASAS users. The extent of the skill decay was small. A group of 21 individuals who received training two months prior to the test scored 90% on a proficiency test, as compared to two individuals who scored an average of 78% one year following training. None of the individuals had any intervening training.

A more controlled examination of decay over a no practice retention interval comes from an experiment by Sanders (1999). In that research project, he examined overlay and report skill decay after a 30 day retention interval. Overlay skills involved creating and sending a graphical map overlay and report skills involved sending text only messages. Participants were trained on these two types of tasks using the Inter-vehicular information system (IVIS), a vehicle mounted digital system that pre-dates FBCB2. In this experiment there was a 52% drop in overlay task proficiency and a 23% drop in report task proficiency.

Clearly, more research is needed in this area to document baseline rates of skill decay over time in current Army ABCS systems. Such information would benefit leaders who must know the frequency and type of training needed to sustain proficiency.

Level of original learning. Level of original learning has been identified as the procedural variable that best predicts skill retention (Hagman & Rose, 1983; Schendel et al., 1978). When it is manipulated experimentally, the two levels of learning most often studied are proficiency training (i.e., training until the task can be completed without error at least once) and mastery training or overlearning (i.e. training some number of additional trials beyond proficiency, e.g., Wisher, Sabol, Ellis, & Ellis, 1999).

In some research reports, where level of original learning varies on a continuous scale (i.e., it is simply measured after training), the degree of original learning predicts a high level of variability in performance at recall. In an experiment on three dimensional flight control, level of original learning was highly correlated with performance at recall (r = .80 to .98) (Fleishman & Parker, 1962). An often cited body of work is that by Bahrick (1979) in which the recall of foreign language (Spanish) words was assessed in a cross sectional experiment. Bahrick found that level of original learning predicted recall as long as 50 years following original learning. Although this type of learning (memorization of vocabulary words) is a fairly simple cognitive skill, research on motor tasks also shows persistent effects of original learning for at least two years (Schendel et al., 1978).

The degree of original learning typically does not change the decay rate so the performance differences present in a baseline test of acquisition are of roughly the same magnitude at recall (Wells & Hagman, 1989; Wisher et al., 1994).Wisher et al. (1994) state in their introduction that a higher level of learning at baseline leads to "slower decay" (p. 3). This choice of words implies that the rate of decay is slower when initial proficiency is higher, but the references cited in support of this statement do not support that conclusion (e.g., Elliott & Wisher, 1993) nor do the results of the Wisher et al (1994) report. It is possible that these authors meant that the skills of individuals trained to a high level of proficiency would take longer to decay to a point when refresher training would be necessary, but this has nothing to do with decay rate.

One notable exception to the rule that differences between proficiency and mastery training groups at initial training are preserved until the retention test is a report by Rose et al. (1985). In this research, performance on 22 cannon crewman tasks was measured in new recruits going through one station unit training. The baseline measure was taken immediately after training was complete and a series of retention tests were given approximately two, four, and six months later. Half of the 145 participants trained to proficiency, that is they trained until they completed a given task without error. The other half received mastery training, that is they continued their training until they had completed a given task two more times without error. At the first retention test, a higher percentage of the mastery group received a "go" on only 4 of the 22 tasks tested. The two groups differed on only one of the 22 tasks at the second and third retention times. The results were much the same when looking at either the number of steps per task completed or the time to complete the task. Furthermore, a regression analysis of the data found that mastery training explained only a small percentage of the total variance (Rose, Czarnolewski et al., 1985).

In a meta-analysis of skill decay by Arthur et al. (1998), 189 variables from 53 articles were analyzed. The authors had expected that level of initial learning would prove to be one of the variables that best predicted skill retention. They found that degree of overlearning only accounted for about 17% of the variability in the retention scores. The reason for this weak effect, the authors suspected, was that only 30 reports examined degree of overlearning and the range of overlearning reported was limited. In essence, Rose et al. (1985) offered the same explanation for their findings. They said that the difference in initial training between their proficient and mastery groups may not have been big enough to produce any lasting effects on recall.

Level of learning and digital skills. Although no reports of the effect of overlearning on digital skill decay could be found, one report on digital skills examined decay as a function of original learning. In his experiment of IVIS skill decay (1999), Sanders compared performance in individuals who successfully completed three of four overlay tasks and two of three of the report tasks to those who completed fewer. About 80% of the sample reached the criterion. The results indicated that those who reached criterion performed significantly better than those who did not after a 30-day retention interval. It should be noted, however, that those who did not reach criterion did not show any skill decay owing to a floor effect in their skill level at both time points.

With the possible exception of the Sanders report just described, there is virtually no data on the impact of level of original learning on the retention of digital skills. Questions about the amount of additional training needed to produce an effect on retention or the duration of such effects remain unanswered. Future research in this area is needed so that the benefits of mastery training can be weighed against the cost of that training.

The spacing of training trials. Another procedural variable that affects recall is the distribution of training trials. Laboratory and field research have shown that for verbal learning, as the training trials get closer together (massed practice), the slope of the learning curve increases, but as the time between training trials increase to a point (distributed practice), performance on retention tests improves (for review see Fendrich et al., 1988; Wells & Hagman, 1989). In other words, the spacing of training trials alters the *slope* of the learning and decay curves. This means that two groups could be trained to an equal level of proficiency, one with spaced trials and one with massed trials, but over time the one with spaced trials would show less skill decay. This means that final performance after acquisition is not the only predictor of performance at recall for verbal tasks.

The effect of the spacing of training trials on learning perceptual motor tasks is less clear. Some authors find that the retention of these types of tasks is comparable whether training is massed or distributed (Fendrich et al., 1988; Wells & Hagman, 1989). This is true for both discrete tasks such as assembling a weapon or for continuous tasks like tracking a target (Schendel & Hagman, 1980). Still other research has found some benefit for spacing practice for motor skills. In an investigation of fuel and electrical repairers who were trained to test electrical alternators using either massed or spaced training, the spaced training group made 40% fewer errors and took half the time to complete the tasks (Hagman, 1980b).

In contrast to verbal learning, researchers generally do not believe that spacing of motor tasks improves learning *per se* but rather the massing of practice on motor tasks leads to fatigue or boredom and lapses in attention that undermine the benefit of training repetition (for review see: Schendel et al., 1978). Because of this possibility, researchers state that spaced practice may be beneficial for motor tasks that are dangerous (i.e., fatigue or a lapse in attention could cause injury) or where Soldiers are not highly motivated (Hagman & Rose, 1983; Wells & Hagman, 1989).

Schedules of training and digital skills. No articles were found for this review that examined the effect of different schedules of training on the retention of computer/digital skills. Given that digital skills are more of a cognitive than a purely motor skill, it is likely that spacing training would provide some improvement in skill retention but the potential gains may not be enough to outweigh the difficulty of scheduling training in a distributed vs. a block schedule. Furthermore, continued training at the unit would, by its nature, be distributed making the distribution of initial training less important. For these reasons, it is probably not fruitful to examine this factor in future work on digital skill retention.

Testing in conjunction with training. In a series of experiments, Hagman (1980a, 1980b) examined the influence different schedules of testing in conjunction with training on a motor task. Participants were asked to move a slider along a track to a preset stop (presentation trial). For the recall test, the stop was removed and participants had to move the slider to where they thought it had been located (testing trial). Participants were given three blocks of six trials and then had two final testing trials 3 min. and 24 hours after the last trial in the third block. Participants were in one of three training schedules: standard - training trials alternated with testing trials; test - one presentation trial was followed by five testing trials for each block of six trials; and presentation - five presentation trials were followed by one testing trial for each block of six trials. As with distributed practice, the group that performed worst during acquisition (test) performed better than the other two groups on recall. The standard group only performed better than the presentation group. Testing in conjunction with training therefore affects decay rate, much as distributed practice does.

Testing in conjunction with training of digital skills. No reports were found which examined testing in conjunction with training for digital skills, but this is an area where research might be fruitful. Digital operator training in the Army tends to be a sequence of guided demonstrations (analogous to Hagman's presentation trials) each followed by a practical exercise (analogous to Hagman's testing trials). This matches the standard condition of the Hagman experiments described in the previous paragraph. Observations of Army digital training by the author have also shown that the practical exercises often turn into guided demos if the students have lots of questions. This makes the training experience more like the demonstration condition (the one showing the worst recall) used by Hagman. To maximize the training benefit, Hagman's findings suggest that multiple practical exercises should follow each demonstration with very little assistance given to the students.

Context interference. Like distributed practice and testing in conjunction with training, a random training schedule *decreases* the acquisition curve relative to a blocked schedule. This benefit of a random schedule is often referred to as the contextual interference effect (for review see Fendrich et al., 1988; Lee & Simon, 2004). The term context in this usage does not refer to the physical environment but rather the context of tasks being performed. A typical experimental protocol would involve training on several tasks. One group would be given a block of training trials on one task and then on the next task and so on (blocked schedule). Another group would be given the same total number of training trials on each task but any trial could be for any of the tasks (random schedule). Generally, the blocked schedule results in better performance during acquisition but the random schedule produces better performance during recall or transfer to a new task (Lee & Simon, 2004). This phenomenon led Battig (1979) to propose the intra-task interference principle of memory. This principle states that greater interference at the time of learning produces higher levels of subsequent retention and transfer. In a review of this literature, Fendrich et al. (1988) found that the context interference effect results in about a 30% - 50% increase in recall performance over a non-interference condition. Thus, the context interference effect slows the rate of decay.

One example of intra-task interference improving a skill of military service members is the report by Hagman (1980b) in which electrical repairers were trained to test alternators. In one condition, the participants were trained to use multiple but similar sets of equipment to do the testing. Interestingly, this condition had no effect on retention but did benefit performance on a transfer task.

Contextual interference and digital skills. There are no examples of research on the context interference effect for digital skills but given the sizable increase in recall performance produced by contextual interference, it would be worth pursuing research in this area. During typical new equipment training (NET) for digital systems, it is common for instructors to progress from one task to the next allowing students time to practice each task in sequence. This is very much like the blocked training condition described earlier. An alternative approach might be to cover a group of tasks and then have the students do a series of practice trials in which the sequence of to-be-practiced tasks is randomized. This could be followed by training on another group of tasks followed by randomized PEs.

A critical principle that emerges from the research reviewed in the last three subsections is that performance during acquisition does not necessarily predict performance during recall. In fact, as the intra-task interference principle indicates, factors that impair performance during acquisition may be the same factors that enhance performance during recall. For this reason, research designed to improve digital skill training (or any other skill) should not rely on a single measure of performance taken immediately after training as the basis for selecting the most effective training technique.

Individual Variables

There are two individual variables that have been investigated with regard to skill decay: aptitude and relevant knowledge. Aptitude or ability can be measured by a variety of techniques including intelligence tests or sub-scales of the Armed Services Vocational Aptitude Battery (ASVAB) such as the Armed Forces Qualification Test (AFQT). Regardless of the measure, all show essentially the same result; individuals with higher ability levels require less time to learn than individuals of lower ability (Adams, 1987; Schendel et al., 1978). In terms of retention, most research finds that aptitude does not affect the rate at which skills are lost, but because individuals with a higher aptitude typically reach a higher level of proficiency following a given training period, those differences in initial proficiency are preserved at the time of the retention test (Hagman & Rose, 1983; Schendel et al., 1978; Wisher et al., 1999).

Several reports indicate that scores on the AFQT predict differences in acquisition and retention measured by written (i.e., cognitive) tests better than hands-on (i.e., motor) tests (Henik, Brainin, Ze'evi, & Schwarz, 1999; Rose, Czarnolewski et al., 1985; Wisher et al., 1994; Wisher, Sabol, Sukenik, & Kern, 1991). In one investigation, primarily examining retention, performance on 18 infantryman tasks was assessed two months following acquisition training (Rose, Czarnolewski et al., 1985). Both mental and hands-on tasks were measured. ASVAB scores predicted performance on mental tasks far better than they predicted performance on hands-on tasks.

In the report by Henik et al. (1999), Israeli Defense Force (IDF) operators of tube-launched, optically-tracked, wire-guided (TOW) missile and M-47 Dragon missile systems were examined. Both a written knowledge test and a hands-on (proportion of hits in a simulator) test were administered 18 months following training. A measure of aptitude (DAPAR, the IDF equivalent of the AFQT) and the score on the qualification test they took following their initial training as well as measures of verbal and visual memory were entered into a regression equation to predict retention test scores. The DAPAR score predicted the written test score but not the hands-on test score, whereas the qualification test score was the only significant predictor of the hands-on test score.

The research by Wisher et al. (1994) looked at how well AFQT scores could predict acquisition of combat engineer skills. In this research IRR Soldiers were examined after a five day rapid train-up exercise. Soldiers were divided into an initial entry training (IET) group who only had minimal training on these tasks when they were on active duty and a prior service group who had completed at least one tour as combat engineers. On a knowledge test of combat engineer tasks, Soldiers who scored above the AFQT median did better than those scoring below the median. Interestingly, for the hands-on test, the AFQT scores made a difference only for the IET (less trained) group. For those in the prior service group, the AFQT did not make a significant difference. This suggests that as hands-on tasks are better learned, they may depend less on verbal and mental ability.

In another investigation of acquisition by Wisher et al. (1991), IRR Soldiers from a wide range of specializations were assessed on written and hands on tests following a rapid train up. Scores on the Soldier qualification test (SQT) significantly correlated with four of five post-training written tests but only one of five post-training hands on tests. The AFQT correlated significantly with one of five written and one of five hands-on tests. Even when the SQT, AFQT, pay grade, and time out of service were included in a multiple regression they accounted for only an average of about 2% of the variability in the scores on any given hands on test, but they accounted for an average of 20% of the variance for the written test scores.

Mental aptitude therefore appears to be a weak predictor of acquisition which can predict differences in proficiency following training. Mental aptitude tests are better at predicting acquisition and retention of cognitive and written tasks than purely motor tasks. Although research generally shows that skill decay rate is independent of aptitude, there is some evidence that lower ability learners forget more abstract and theoretical material than high ability learners (Arthur et al., 1998).

Aptitude and digital skills. There are virtually no data on the influence of aptitude on the acquisition of digital skills. Although some researchers (Davis & Yi, 2004; Simon & Werner, 1996) measure aptitude when studying digital skill training, it is only to filter out the confounding the effects of aptitude on the independent variable. As will be shown later in this report, a more relevant issue may be the interaction between ability and method of training. Individuals of lower ability probably respond better to instructor led training than self-guided training.

Background knowledge. In addition to training on a given ABCS component, it may be beneficial for Army learners to have knowledge about computers, computer software, computer networks, as well as a knowledge of map symbols, graphic control measures, and even global positioning system technology. In a paper describing ways to train officers to exploit MCS, Leyden (2002) proposes that officers receive training in computer network configuration, system components, and equipment requirements.

A few reports have confirmed the benefits of background knowledge when learning new digital systems. In the investigation of recall performance of IVIS skills (Sanders, 1999) declarative knowledge about the IVIS system was significantly correlated with total successful overlay trials during training. In addition, participants' use of computers was significantly and positively correlated with their 30-day recall performance. In another investigation, participants' knowledge of a mobile subscriber network was a significant predictor of their retention of complex procedural skills needed to operate the network (Wisher et al., 1994). Finally, in the experiment by Dyer and Salter (2001) on the training of digital map interface skills, background knowledge of military map symbols and computer skills predicted scores on the final test in a regression analysis. One paper reported that the beneficial effects of a general knowledge of computer networks was not in learning to operate the software (Elliott, Sanders, & Quinkert, 1996) but rather in learning how to troubleshoot when the system failed.

A web-based training program was developed to improve the skills of AFATDS operators by increasing their background in computer systems (Hess, Alliger, Linegang, Meischer, & Garrity, 2003). This product, called the Learning Skills Bridge Learning Accelerator, was found to significantly improve performance on a test of AFATDS knowledge. The learning accelerator was intended to help AFATDS operators more quickly adapt to changes in software and hardware design. Unfortunately the learning accelerator was only tested on a single group of subjects, all of whom received the training. Without a control group, it was impossible to determine the effects of the learning accelerator over a no-training condition. Furthermore, the test diagnosed their knowledge of AFATDS, but did not examine how well individuals could apply their knowledge to a new system.

In summary several reports indicate that background knowledge has an impact on the training and retention of digital skills. This can easily be explained from the perspective of cognitive psychology as it would be expected that individuals with greater background knowledge can more easily organize and encode the new information since it is more likely to fit into already existing mental schemas. This suggests that classroom time spent explaining the systems may facilitate the learning of procedural skills. Future research is needed to determine how much or what types of background knowledge will benefit students learning or retraining on Army digital systems.

Task Variables

Certain characteristics of the task being trained predict the rate at which the task is likely to decay (Rose, Czarnolewski et al., 1985). In fact, the correlation between actual and predicted retention performance in this research is in the range of $r = 0.90$. This is not to say that training or individual variables only account for a small proportion of variability, but rather when all other variables are held constant, task variables are good predictors of skill decay. An advantage of using task variables, rather than procedural or individual variables, to predict skill decay is that a single subject matter expert can make the prediction. Using a variable like the level of initial training to predict skill decay requires a time consuming and expensive data collection and analysis effort.

To better understand the task variables that affect skill decay, it is necessary to understand the distinctions between different types of memory. Cognitive psychologists have long recognized a distinction between declarative and procedural memory. Declarative memory is comprised of explicit facts and information, whereas procedural memory is a memory for how to do things. In addition to the logical distinction between these types of memory, there is a large body of neuropsychological evidence to suggest that these types of memory are mediated by independent brain systems (see Gabrieli, 1998 for review).

Although the ability to remember facts (declarative memory) is not a skill, it is a necessary ability for the performance of other skills. Declarative memory has been shown to be fairly resistant to decay as evidenced in the research of Bahrick (1979) on the recall of foreign language vocabulary words. Other evidence of the persistence of declarative memory comes from the experiment of IDF missile operators, in which knowledge about procedural skills did not decline until after 12 months (Henik et al., 1999). Similarly, in a review by Wisher et al. (1999), memory for decision skills and job knowledge showed relatively minimal decay (generally less than 20% loss) over a two year period.

Most skills fall under the heading of procedural rather than declarative memory, and it is commonly believed that procedural skills are very resistant to decay. The use of the phrase, "It's like learning how to ride a bike," to refer to a task that is never forgotten once learned shows how prevalent the belief is. This belief, however, is only partially true.

By the late 1950s, psychologists realized that not all procedural tasks were resistant to decay. Continuous tasks with no beginning or end (tasks like riding a bike, sometimes referred to as open-loop tasks) were, in accordance with the above phrase, resistant to forgetting. On the other hand, discrete procedural tasks (tasks that have a discreet beginning and end, sometimes referred to as closed-loop tasks) tended to be easily forgotten (Adams, 1987).

Research in the field of pilot skill decay clearly shows the different rates of perishability for open- and closed-loop skills. For example, there is almost no decay of continuous flying skills (those needed to maneuver the aircraft) over months or years but discreet pilot procedures (i.e., engine startup and shutdown) decay to unsafe levels within a matter of weeks or months without practice (Schendel et al., 1978).

Given the relatively greater decay rate for discrete procedural skills, most research has been focused on improving retention of this type of skill (Rose, Czarnolewski et al., 1985; Sanders, 1999; Shields, Goldberg, & Dressel, 1979). With regard to military skills, a paper by Rose et al. (1985) has shown that for discreet procedural skills, task complexity, the demands of the task, the availability of job aids, and the presence of stress or a time limit all affect decay rate. These factors have been reviewed in detail in several reports (Hagman & Rose, 1983; Wells & Hagman, 1989; Wisher et al., 1999) and so will be covered briefly here.

Task complexity has to do with how many steps there are in a task, whether the steps must be performed in a specified sequence, and whether there is built in feedback that indicates correct performance of the task. The clearest demonstration of the impact of the number of steps on skill retention was a report of field artillery tasks (Shields et al., 1979). In this survey, the percent of the sample that could perform the task after a one year retention interval declined significantly as a function of the number of steps in the task with only 20% of the sample still able to perform tasks with 12 or more steps.

Task demands include cognitive, knowledge, or execution demands. Tasks may require individuals to recall definitions, names, or locations. Tasks that require the recall of fewer than 8 items are remembered well but tasks that require the recall of more than 8 items suffer rapid decay (Wisher et al., 1999). In general, the greater the physical task demands the faster the decay, although surprisingly tasks that require only simple motor control such as hammering a nail decay faster than tasks that require moderate precision (Rose, Radtke, Shettel, & Hagman, 1985a).

The availability of job and memory aids will generally aid in recall (Rose, Czarnolewski et al., 1985), but not if the task can be easily performed from rote memory. In an experiment examining the benefits of a mnemonic for installing the M14 antipersonnel mine, no benefit was found. Participants reported that the task was easy to recall thus high performance in the control group negated any benefits of the memory aid (Hagman & Rose, 1983).

A predictive model for military skill decay. In the early 1980's ARI undertook an effort to develop an empirically based model for predicting skill decay (Rose, Czarnolewski et al., 1985; Rose, Radtke et al., 1985a). The model was designed so that unit leaders could estimate how quickly skill decay would occur for any given skill and subsequently how often refresher training would be needed to maintain the level of proficiency desired by the leader.

Predictions of skill decay are made using the Users Decision Aid or UDA which takes ratings of task variables and uses them to predict skill decay (Rose, Radtke et al., 1985a; Rose, Radtke, Shettel, & Hagman, 1985b). The result is an estimate of the percentage of individuals who will be proficient on a given task after a specified retention interval assuming everyone in the group is initially trained to proficiency.

Ratings used to generate estimates by the UDA are made by subject matter experts familiar with the task to be trained. The questions pertain to the factors described above (i.e., the number of steps required; ratings of the mental and motor control demands, etc.). Data show that expert raters will generate very consistent ratings (correlations greater than 0.90).

The Refined Users Decision Aid (UDA) consists of 10 questions, and the strongest predictors deal with mental challenge (Rose, Czarnolewski et al., 1985). The questions that carry most of the predictive weight have to do with the presence of feedback from each step as to whether it was performed correctly, the mental challenge of the task, the number of facts that must be recalled and the difficulty of recalling them. It is important to note that the UDA assumes all individuals start at 100% proficiency (i.e., there are no differences in baseline performance) and it does not take into account the effects of particular training methodologies.

The UDA produced very accurate estimates of skill decay when it was validated using measures of performance of field artillery tasks over time with correlations between the actual and predicted performance generally in the $r = .90$ range (Rose, Czarnolewski et al., 1985). By comparison, the predictions made by the UDA were superior to those based on ASFAB Field artillery subtest scores (correlations in the $r = .30$ range) but comparable to predictions based on baseline proficiency measures. The authors pointed out that using the UDA has a major advantage over baseline proficiency measures in that the collection, and analysis of those measures is complex and time-consuming (Rose, Czarnolewski et al., 1985).

Use of the UDA to predict the decay of digital skills. As reviewed by Wisher et al. (1999), the UDA has been applied in a number of military specialties including vehicle mechanics, radio operators, quartermasters, combat engineers, field medics, and air defense missile crews. More recently, it has been used to predict the decay of IVIS skills (Sanders, 1999). The IVIS system is no longer in use but it performed many of the same functions as the current FBCB2 system allowing the Abrams Tank crew to see their location displayed on a digital map and to send and receive overlay and message information.

Participants in the IVIS skill retention experiment performed a series of overlay tasks involving creating and sending overlays and a series of communications tasks involving creating and sending messages. Skill decay was measured 30 days following acquisition training. In general, the UDA under-predicted skill decay. The overlay skills were predicted to be retained at 67% but in fact were retained at 48% (difference not significant). The report skills were predicted to be retained at 92% but in fact were retained at 77% (significant difference). Some of the reason for the discrepancies between actual and predicted skill levels may have had to do with software eccentricities (i.e., clicking a "send" button does not send an overlay) or training shortfalls (i.e., training was not provided to correct errors in data entry). This suggests that some modification of the UDA may be needed to better predict digital skill decay (Sanders, 1999).

Since the UDA is based only on an analysis of the task, it would be possible to use this tool to estimate the decay rate of tasks on various ABCS systems. Even if the absolute levels of decay are not estimated with 100% accuracy, knowing the relative levels of decay would still be valuable. The identification of the most and least perishable digital skills would help training planners focus training on the skills that need it the most.

Digital Skill Assessment

The assessment of digital skills is important not only for research but also for units to know how often and what digital skills to train (Moses, 2001). Ideally, assessment tools should be reliable, valid, and easy to use. These goals are not daunting when developing a tool to assess the skills of an individual operator; however, it can be a challenge to achieve these goals when developing a tool to evaluate the performance of a command group employing multiple systems.

Tests of system operators typically cover functions an operator would commonly use (e.g. creating and saving messages or overlays, displaying information on a map, basic troubleshooting, etc.). These kinds of skills are discrete, multi-step procedures that can be easily measured. The assessment of the collective employment of digital systems, on the other hand, is far more challenging. Employment of these systems involves following unit standing operating procedure (SOP) on how to process and distribute information across and within echelons. Measuring the simultaneous behavior of a group of individuals as they share and process information electronically presents its own unique set of challenges.

There are two areas of research focused on the measurement of digital skills. The first is a set of reports on the development and use of digital proficiency measurement guides for current ABCS systems and the second is a series of experiments examining human performance using a hypothetical Future Combat System (FCS). These two areas are briefly reviewed below.

Digital Proficiency Guides

During training, instructors or observer/controllers (O/Cs) are typically responsible for assessing performance. This feedback is a critical part of the training experience and instructors and O/Cs need proficiency measurement tools to provide valid and reliable feedback to the units that are training. Towards that end, researchers have developed digital proficiency measures for units doing individual or collective digital training (Barnett, Meliza, & McCluskey, 2001; Leibrecht, Lockaby, & Meliza, 2003a, 2003b; Leibrecht, Lockaby, Perrault, & Meliza, 2004a).

The main purpose of the digital proficiency guides is to reduce the workload of trainers and O/Cs to a manageable level by focusing them on high payoff measurement targets (Leibrecht et al., 2003a). Much of this is accomplished by organizing all of these high payoff measurement targets into a logical taxonomy. For example two of the guides are designed for FBCB2, the *FBCB2 Leader's Primer* and the *FBCB2 Exploitation Tool* (both described in Leibrecht et al., 2003b). The Leader's Primer takes a higher level look at employment of FBCB2 than the

Exploitation Tool. The Leader's Primer breaks down FBCB2 tasks into 5 categories (digital basics, battlefield visualization, mission planning and preparation, tactical information exchange, and force mobility and maneuver). The Exploitation Tool breaks down FBCB2 tasks into 9 "skill groups" (perform precombat checks, disseminate and manage messages and graphics, plan and execute movements, apply situational understanding in maneuver decisions, conduct collaborative planning, support logistical operations, control indirect fires, avoid fratricide, and employ filter settings). Under these categories are specific performance goals or keys to success. Along with these are tips on how to know if the goals or keys are being met and why they are important (Leibrecht et al., 2003b).

For the digital tactical operations center (TOC) the command and control center for a battalion or brigade, *The Digital TOC Integration Guide* was developed (Leibrecht, Lockaby, Perrault, & Meliza, 2004b). This guide is comparable to the FBCB2 Exploitation Tool in that it is aimed at providing detailed feedback to the staff members. The Integration Guide breaks tasks down into three groups it calls "Integration Skills" (establish and manage the common operating picture (COP), manage digital info, and avoid fratricide). Each Integration skill is further broken down by battlefield operating system (BOS) and associated staff sections and each of these lists multiple "responsibilities" (e.g., save MCS overlays as a .mgc file). The guide also describes ways for O/Cs to confirm that these responsibilities are accomplished (e.g., by asking questions or observing the systems).

Another set of tools are the *Battle Staff Proficiency Level Tables* that allow unit leaders or O/Cs to rate various staff sections at a low, medium, or high level of proficiency across a variety of tasks (Leibrecht et al., 2004a). These Proficiency Tables are designed to be analogous to the FBCB2 Leaders Primer in that they provide a higher level look at performance in the digital TOC.

Finally, a set of quick assessment guides (QAGs) were developed for both FBCB2 and digital TOC operations. They contain (40 - 50 yes/no questions) and they are specifically designed to allow leaders to quickly determine whether their unit is performing at a basic, medium, or high level of proficiency (Leibrecht et al., 2006). Once leaders determine the level their unit is performing at, there are basic, medium, and high training guides that help leaders determine specific areas their unit needs training in to achieve a Basic/Medium/High level of proficiency. The training guides are categorized by "Skill" (plan, prepare, execute) and each skill is further subdivided by "Skill group" (channel info, manage info, assess info, exploit info). Each skill group is further subdivided by staff section.

A strategy to tailor training for a particular unit using all of these proficiency guides is provided by Leibrecht et al. (in preparation). The general strategy is to start with tools that provide a high level assessment and then gradually move to tools that provide a finer grained analysis of skills. More specifically, the first step of the strategy is to use the Quick Assessment Guides to determine unit proficiency at a gross level (basic, medium, or high). From this assessment the unit would develop a training plan using the basic, medium, and high training guides and execute it using the FBCB2 Exploitation Tool or Digital TOC Integration Guide to provide training feedback (Leibrecht et al., 2006)

All of these proficiency measures were developed based on guidance from leaders of digitized units in both the 1st Cavalry Division and the 4th Infantry Division (Leibrecht et al., 2003a; Leibrecht et al., 2004b). They are all currently available from the Battle Command Training Center's Digital Reference Center (https://bctc.army.mil) for downloading and they are among the most frequently downloaded files among a large library of files. Although to date, no formal validation efforts with training units have been published, the informal feedback from units using these guides for training has been consistently positive. Furthermore, validation of these guides using selected subject matter experts has generally confirmed that the items on the guides accurately measure unit proficiency (Leibrecht et al., 2006).

Assessing Collective Digital Skills in Future Combat Systems

Digital command and control in Future Combat Systems is anticipated to build on current ABCS technology. To investigate how such a system would affect battle command and to explore the training requirements for future C2 systems, a series of experiments were conducted using a notional FCS command and control mock-up vehicle (Carnahan, Lickteig, Sanders, Durlach, & Lussier, 2004; Lickteig, Sanders, Durlach, & Carnahan, 2004; Lickteig, Sanders, Durlach, Lussier, & Carnahan, 2003). The basic experimental design allowed participants to spend three days training on the prototype FCS. This was followed by six days of experimental trials in which participants engaged in a simulator driven exercise. A key feature of the trials was that many aspects of the mission remained constant (goals and terrain) whereas some aspects varied (number of enemy units, restrictions of certain supporting assets, presence of civilians on the battlefield). The goal was to give the unit practice with some aspects of the mission across trials while varying the complexity of the trials thereby minimizing the training requirements across trials.

The primary participants were four active duty Lieutenant Colonels and a Major who served as an alternate. Four main experiments were run to examine the performance of an expert group and a fifth experiment with Army Cadets was run to examine differences between novice and expert command groups.

The primary dependent measures were verbal communications, human computer interactions, and self-report subjective measures. Verbal communications were broken down by duty position (commander, battlespace, information, and effects) and function (plan, move, see, strike, battle damage assessment, and other). Human computer interactions were recorded by videotaping the computer screens of the FCS C2 systems and then grouping interactions in one of four categories (plan, see, move, strike). Each category was further subdivided into four or five sub-categories. The subjective results included workload, performance, and effectiveness ratings (Lickteig, Sanders, Lussier et al., 2004).

Verbal interaction was found to be almost continuous (93% of the time) even though all had access to a rapidly updated and accurate common operational picture (COP). Most of the verbalizations were made by the commander (about 55% vs. 18% by others averaged across experiments). Most of the verbalizations were in the "see" function (30%) and the "strike" function (22%). Interestingly, the commander, who spent most of his time verbalizing, had the lowest frequency of human computer interactions (150 interactions per trial vs. 350 interactions for the battle managers). Workload ratings varied as a function of trial complexity as expected. In addition, the workload ratings for the information manager were consistently the highest. Interestingly, across experiments software changes were made to automate and theoretically reduce the workload of the command group. Sometimes these were successful and sometimes they had the opposite effect because they led to the expectation that the command group could accomplish the mission in less time (Lickteig, Sanders, Lussier et al., 2004).

The comparison of novice (cadets) and experienced battle commanders revealed some interesting differences. The novices spent less time collaborating and hastily deployed their forces to find and destroy the enemy. For example they performed about half as many friendly and enemy queries as did experts (to perform a query, which produced information about the unit, the participant simply moved his cursor over the unit. Upon doing this, the unit's size, direction, rate of movement, and status were all displayed). The experts on the other hand were more collaborative and spent more time building an accurate and complete understanding of the battlefield before committing to action (Carnahan et al., 2004).

Training across all experiments was evaluated primarily by use of questions directed at the participants. The first two days of training in each experiment were spent learning individual operator skills and collective mission training was done on the last day of training only. There was also learning taking place across experiments, as the same command group was used in all four main experiments. Even with the operational experience that the four expert participants possessed and their experience participating in over 40 experimental trials, they still felt that they needed more hands on training focused on employing the systems. One thing they specifically wanted was to be able to run through a couple of missions to help reestablish and refine SOPs. Coincidentally, the desire for more hands-on training and scenario-based collective training are the most commonly requested types of training by current users of the Army's ABCS systems (Schaab et al., 2004).

The experimenters also noted that greater automation required more training on the logic of the automated process. They cited an example of an automated unmanned aerial reconnaissance vehicle (UAV) that tended to wander off course and get destroyed in one of the experiments. It was determined that this was due to the participants misunderstanding the rules that determined the path of the UAV. The authors concluded that the training challenge associated with more advanced and automated C2 systems could not be underestimated (Lickteig, Sanders, Durlach et al., 2004; Lickteig et al., 2003), a conclusion that was also reached by Schaab and Moses (2001).

It is difficult to directly apply research on the training of a prototype FCS to current ABCS systems but the research methodology used in these reports could readily be adapted to the assessment the collective employment of current ABCS systems. The lessons learned by these investigators as to how best to capture the complex interactions of a command group using digital C2 systems should provide considerable savings of effort for researchers designing experiments to investigate training and retention of these ABCS skills.

Conclusions and Future Directions

Digital Skill Decay. Research on individual digital operator skills has highlighted a number of factors likely to improve retention of digital skills and has pointed out directions for future research. Despite the widespread belief that digital skills are perishable, little empirical data backs this up. The experiment by Sanders (1999) is the only research to date that documents decay rate (he observed from 30% - 50% decay), but it is based on a system no longer in use in the Army. A better understanding of the perishability of digital skills organized by system and task is badly needed to help target training where it is needed.

Experimentally determining the decay rate of digital skills on all ABCS systems would be prohibitively expensive and time consuming. A much wiser approach would be to use the UDA to predict skill decay on key procedures on any given system and then validate only a selected subset of those procedures.

In addition to a more complete knowledge of the decay rates of digital skills, there have been no reports looking at how long it takes to retrain skills once they are forgotten. Anecdotal reports from active duty FBCB2 users indicate that although their digital skills are easily forgotten, they feel confident that with only an hour or two of self-guided exploration, they could restore their proficiency levels. A better empirical understanding of the frequency, type, and amount of refresher training needed to sustain digital skills would be valuable to units who need to work this into their training schedules.

Research on overlearning and the spacing of training trials has shown that these two factors can improve skill retention but it is not clear that research in these areas would be particularly fruitful. One reason is that the size of the effect of these factors is variable and another reason is that even if they were found to be beneficial, implementing these practices may not be practical.

Digital Skill Training. The incorporation of intratask interference and training principles such as guided exploration into operator training is likely to benefit the design of operator training courses. The research by Schaab and Dressel (2001) is an excellent example of how a constructivist approach can benefit Soldiers doing digital training. One caveat of this experiment is that it was done on Military Intelligence Analysts, a group that has been selected for high mental aptitude. Given research showing that individuals with lower aptitudes or abilities tend to do better in more structured learning environments (e.g., Baldwin et al., 1976; Dyer et al., 2005) additional research on the use of these techniques on a less selected student population should be done to determine whether this approach is beneficial for the training of all digital systems.

Research on training collective digital skills is an area that would be of tremendous benefit to the Army. Digital ABCS systems are still relatively new to many Army units and there is not a large base of experienced leaders to guide units as they learn to employ these systems (J. E. Clark, 2005). Additionally, research on how to best manage information coming across a digital network and how to cope with information overload would be beneficial. Research on these topics would help to define "what right looks like" for units learning to use these systems so that they can model their behavior accordingly. All future research on the training of digital skills should be careful to avoid relying on end-of-training exams as measure of the best training approach. As reviewed above, the approaches that lead to the fastest acquisition often do not produce the best long-term retention.

Digital Skill Assessment. The set of proficiency guides that have been developed for current Army ABCS (version 6.4) systems reviewed above have been well received by units who've used them, but little work has been done to validate their use in a training environment. If units are employing these guides to make training decisions, it might be possible to gather some data from those units to both validate the guides and refine recommendations for their use.

There are potential benefits for measures of collective system employment in the realm of basic research. For example, virtually nothing is known about how long it takes a novice command group to reach expert status using ABCS systems nor of the training techniques that will most facilitate that learning. The investigation of novice and expert users of the FCS system by Carnahan et al. (2004) is the kind of research that might be done to better understand this process. As stated above, research in this area is challenging and resource intensive meaning that only a limited number of projects will be possible. This is a particular area where collaboration across ARI field units may be needed to pool both resources and expertise.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin, 101*(1), 41-74.

Advanced warfighter experiment focused dispatch final report. (1996). Fort Knox, KY: U.S. Army Armor Center, Mounted Battlespace Battle Lab.

Archer, R., Warwick, W., McDermott, P., & Katz, J. (2003). *Training future force leaders to make decisions using digital information* (ARI-Research Report-1815XA-ARI). Alexandria, VA: U.S. Army Research Institute. ADA419668

Arthur, W., Jr., Bennett, W., Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance, 11*(1), 57-101.

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General, 108*(3), 296-308.

Baldwin, R. D., Cliborn, R. E., & Foskett, R. J. (1976). *The acquisition and retention of visual aircraft recognition skills* (ARI Technical Report 76-A4). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA035863

Barnett, J. S., Meliza, L. L., & McCluskey, M. R. (2001). *Defining digital proficiency measurement targets for U.S. Army units* (ARI Technical Report 1117). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA396899

Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing and human memory* (pp. 23-44). Hillsdale, NJ: Erlbaum.

Carnahan, T. J., Lickteig, C. W., Sanders, W. R., Durlach, P. J., & Lussier, J. W. (2004). *Novice versus expert command groups: Preliminary findings and training implications for future combat systems* (ARI Research Report 1821). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA421871

Carroll, J. M., Mack, R. L., Lewis, C. H., Grischkowsky, N. L., & Robertson, S. R. (1985). Exploring a word processor. *Human-Computer Interaction, 1*, 283-307.

Charney, D., Reder, L., & Kusbit, G. W. (1990). Goal setting and procedure selection in acquiring computer skills: A comparison of tutorials, problem solving, and learner exploration. *Cognition and Instruction, 7*(4), 323-342.

Clark, J. E. (2005). *Solving a command and control system education and training dilemma for the modular force* (A white paper). Washington, D.C.: Army Joint Support Team.

Clark, R., & Wittrock, M. C. (2000). Psychological principles in training. In S. Tobias & J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 51-84). New York: Macmillan.

Czaja, S. J., Hammond, K., Blascovich, J. J., & Swede, H. (1986). Learning to use a word-processing system as a function of training strategy. *Behaviour & Information Technology, 5*(3), 203-216.

Davis, F. D., & Yi, M. Y. (2004). Improving computer skill training: Behavior modeling, symbolic mental rehearsal, and the role of knowledge structures. *Journal of Applied Psychology, 89*(3), 509-523.

Deatz, R. C., & Campbell, C. H. (2001). *Application of cognitive principles in distributed computer-based training* (ARI Research Product 2001-03). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA392929

Dryburgh, H. (2002). Learning computer skills. *Education Quarterly Review, 8*(2), 8 - 18.

Dyer, J. L., & Salter, R. S. (2001). *Working memory and exploration in training the knowledge and skills required by digital systems* (ARI Research Report 1783). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA399507

Dyer, J. L., Singh, H., & Clark, T. L. (2005). *Computer-based approaches for training interactive digital map displays* (ARI Research Report-1842). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA440171

Ebbinghaus, H., Ruger, H. A., & Bussenius, C. E. (1913). *Memory: A contribution to experimental psychology*. New York, NY: Teachers College Press.

Elliott, G. S., Sanders, W. R., & Quinkert, K. A. (1996). *Training in a digitized battalion task force: Lessons learned and implications for future training* (ARI Research Report 1695). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA313293

Elliott, G. S., & Wisher, R. A. (1993). *A review and annotated bibliography of literature relevant to armor skill retention* (ARI-93-03). Alexandria VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA263407

Fendrich, D. W., Healy, A. F., Meiskey, L., Crutcher, R. J., Little, W., & Bourne, L. E., Jr. (1988). *Skill maintenance: Literature review and theoretical analysis* (Air Force Human Resources Laboratory Technical Paper 87-73). Brooks Air Force Base, TX: Air Force Systems Command. ADA198757

Fleishman, E. A., & Parker, J. F., Jr. (1962). Factors in the retention and relearning of perceptual-motor skill. *Journal of Experimental Psychology, 64*(3), 215-226.

Folds, D., Blunt, C., & Stanley, R. (in preparation). *Training for rapid interpretation of voluminous multimodal data* (ARI Research Report). Alexandria, VA: U.S. Army Research Institute for the Behaivoral and Social Sciences.

Frese, M., Brodbeck, F., Heinbokel, T., Mooser, C., & et al. (1991). Errors in training computer skills: On the positive function of errors. *Human-Computer Interaction, 6*(1), 77-93.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49*, 87-115.

Gist, M. E., Rosen, B., & Schwoerer, C. (1988). The influence of training method and trainee age on the acquisition of computer skills. *Personnel Psychology, 41*(2), 255-265.

Gist, M. E., Schwoerer, C., & Rosen, B. (1989). Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology, 74*(6), 884-891.

Hagman, J. D. (1980a). *Effects of presentation- and test-trial training on acquisition and retention of movement* (ARI Technical Report 492). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA100867

Hagman, J. D. (1980b). *Effects of training schedule and equipment variety on retention and transfer of maintenance skill* (ARI Research Report 1309). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA120167

Hagman, J. D., & Rose, A. M. (1983). Retention of military tasks: A review. *Human Factors, 25*(2), 199-213.

Healy, A. F., Meiskey, L., Fendrich, D. W., Crutcher, R. J., & Little, W. (1988). *Skill maintenance: Literature review and theoretical analysis* (AFHRL Technical Paper 87-73). Brooks Air Force Base, Texas: Air Force Human Resources Laboratory. ADA198757

Henik, A., Brainin, E., Ze'evi, V., & Schwarz, D. (1999). *The preservation and the decay of military skills* (ARDSG-R/D 7699-RB-01). Beer-Sheva, Israel: Department of Behavioral Sciences University of the Negev. ADA389399

Hess, K. P., Alliger, G., Linegang, M., Meischer, S., & Garrity, M. J. (2003). *Training adaptability in digital skills: The learning skills bridge (LSB) learning accelerator* (ARI Research Report 1811). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA416238

Johnston, J. C., Leibrecht, B. C., Holder, L. D., Coffey, R. S., & Quinkert, K. A. (2003). *Training for future operations: Digital leaders' transformation insights* (ARI Special Report 53). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA412717

Lee, T. D., & Simon, D. (2004). Contextual interference. In A. M. Williams & N. J. Hodges (Eds.), *Skill acquisition in sport: Research, theory and practice* (pp. 29-44). London, UK: Routledge.

Leibrecht, B. C., Lockaby, K. J., & Meliza, L. L. (2003a). *Exploiting FBCB2 capabilities through realistic feedback* (ARI Research Report 1810). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA415999

Leibrecht, B. C., Lockaby, K. J., & Meliza, L. L. (2003b). *A practical guide for exploiting FBCB2 capabilities* (ARI Research Product 2003-05). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA415997

Leibrecht, B. C., Lockaby, K. J., Perrault, A. M., & Meliza, L. L. (2004a). *Digital proficiency levels for the brigade and battalion battle staff* (ARI Research Report 1826). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA425934

Leibrecht, B. C., Lockaby, K. J., Perrault, A. M., & Meliza, L. L. (2004b). *Measuring digital battle staff proficiency in current and future forces* (ARI Research Report 1825). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Leibrecht, B. C., Lockaby, K. J., Perrault, A. M., Strauss, C. P., & Meliza, L. L. (2006). *Tailored feedback for digital training* (ARI Research Report 1858). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Leibrecht, B. C., Lockaby, K. J., Perrault, A. M., Strauss, C. P., & Meliza, L. L. (in preparation). *Tailored feedback for digital training*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Leyden, R. G. (2002). *Training field grade officers to exploit the maneuver control system* (Monograph). Fort Leavenworth, KS: School of Advanced Military Studies Army Command and General Staff College. ADA403727

Lickteig, C. W., & Emery, C. D. (1994). *Information management performance of future platoon leaders: An initial investigation* (ARI Technical Report 1000). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA282908

Lickteig, C. W., Sanders, W. R., Durlach, P. J., & Carnahan, T. J. (2004). *Future combat systems command and control (FCS C2) human functions assessment: Interim report - experiment 3* (ARI Research Report 1819). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA420960

Lickteig, C. W., Sanders, W. R., Durlach, P. J., Lussier, J. W., & Carnahan, T. J. (2003). *Human performance essential to battle command: Report on four future combat systems command and control (FCS C2) experiments* (ARI Research Report 1812). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA419499

Lickteig, C. W., Sanders, W. R., Lussier, J. W., & Durlach, P. J. (2004). *Measurement of human performance for future combat systems command and control* (Conference Paper).

Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA433110

Lynch, R. P. (2001). *Lessons learned: Commanding a digital brigade combat team* (IDA Paper P-3616). Alexandria, VA: Institute for Defense Analysis-Joint Advanced Warfighting Program. ADA395042

Mayer, R. E. (1975). Information processing variables in learning to solve problems. *Review of Educational Research, 45*(4), 525-541.

Moses, F. L. (2001). *Training challenges for digitization* (ARI Special Report 47). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA395401

Palmiter, S., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction, 8*(3), 193-216.

Pear, T. H. (1927). Skill. *Personnel Journal, 5*, 478-489.

Rose, A. M., Czarnolewski, M. Y., Gragg, F. E., Austin, S. H., & Ford, P. (1985). *Acquisition and retention of Soldiering skills* (ARI Technical Report 671). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA160336

Rose, A. M., Radtke, P. H., Shettel, H. H., & Hagman, J. D. (1985a). *User's manual for predicting military task retention* (ARI Research Product 85-13). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA160416

Rose, A. M., Radtke, P. H., Shettel, H. H., & Hagman, J. D. (1985b). *User's manual for predicting military task retention. Volume 2. Revision* (ARI Research Product 85-26). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA163710

Sanders, W. R. (1999). *Digital procedural skill retention for selected M1A2 tank inter-vehicular information system (IVIS) tasks* (ARI Technical Report 1096). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA368212

Sanders, W. R. (2001). *Cognitive psychology principles for digital systems training* (ARI Research Report 1773). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA391035

Schaab, B. B., & Dressel, J. D. (2001). *Training for adaptibility and transfer on digital systems* (ARI Research Report 1782). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA399409

Schaab, B. B., & Dressel, J. D. (2003). *Training the troops: What today's Soldiers tell us about training for information-age digital competency* (ARI Research Report 1805). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA415297

Schaab, B. B., Dressel, J. D., & Moses, F. L. (2004). *Digital skills training for net-centric operations* (ARI Special Report 58). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA421861

Schaab, B. B., & Moses, F. L. (2001). *Six myths about digital skills training* (ARI Research Report 1774). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA392922

Schendel, J. D., & Hagman, J. D. (1980). *On sustaining procedural skills over prolonged retention intervals* (ARI Research Report 1298). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA120758

Schendel, J. D., Shields, J. L., & Katz, M. S. (1978). *Retention of motor skills. Review* (ARI Technical Paper 313). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA061338

Shields, J. L., Goldberg, S. L., & Dressel, J. D. (1979). *Retention of basic Soldiering skills* (ARI Research Report 1225). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA075412

Simon, S. J., & Werner, J. M. (1996). Computer training through behavior modeling, self-paced, and instructional approaches: A field experiment. *Journal of Applied Psychology, 81*(6), 648-659.

Swinford, P. L. (1997). *Decision making implications of digital information systems for the battalion in combat* (Master's Thesis). Fort Leavenworth, KS: Army Command and General Staff College. ADA331872

Throne, M. H., & Lickteig, C. W. (1997). *Training computer skills for the future battlefield: A review and annotated bibliography* (ARI Research Product 97-16). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA337314

Wells, R., & Hagman, J. D. (1989). *Training procedures for enhancing reserve component learning, retention, and transfer* (ARI Technical Report 860). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA217450

Wisher, R. A., Kern, R. P., Sabol, M. A., & Farr, B. J. (1994). *Reacquisition of skills by combat engineers mobilized from the individual ready reserve* (ARI Research Report 1665). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences. ADA286244

Wisher, R. A., Sabol, M. A., Ellis, J., & Ellis, K. (1999). *Staying sharp: Retention of military knowledge and skills* (ARI Special Report 39). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences. ADA366825

Wisher, R. A., Sabol, M. A., Sukenik, H. K., & Kern, R. P. (1991). *Individual ready reserve (IRR) call-up: Skill decay* (ARI Research Report 1595). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences. ADA239371