The Training and Retention of Selected FBCB2 Operator Skills

Gregory Goodwin, Jean Dyer
U.S. Army Research Institute
Fort Benning, Georgia
gregory.goodwin1@us.army.mil, jean.dyer@us.army.mil

Bruce Leibrecht, Richard Wampler
Northrop Grumman Corp.
Columbus, Georgia
bruce.leibrecht@ngc.com, rich.wampler@ngc.com

ABSTRACT

It is commonly believed that digital skills are highly perishable, yet little empirical data exist to document just how perishable these digital skills are. Skill decay for Force XXI Battle Command Brigade and Below (FBCB2) was investigated with 54 infantry captains who attended a two-day FBCB2 familiarization course. Operator knowledge and skills were measured immediately after the course and again eight weeks later. During the eight-week retention interval, none of the participants had an opportunity to practice with an FBCB2 system. At baseline, individuals did not do as well on the knowledge test (avg. score 40%) as they did on the hands-on test (avg. score 72%). After the eight-week retention interval, there was no significant overall decline in performance on the knowledge test, but this may have been due to a floor effect. On the other hand, there was a small (10%) but significant forgetting of operator skills. Decay of knowledge and skills was not uniform across individual items. On the knowledge test, forgetting was most acute for a specific fact (the packet mode message size limit in bytes – 45% decline). On the operator skill test, auto-centering the icon on the map, creating a route, and creating an address group all showed significant declines of 20% or more while creating and sending free-text messages and using the line-of-sight tool showed virtually no decay. Measures of training, experience, and knowledge were all examined as possible mediators of skill decay but were not found to have a very large impact on proficiency scores. In summary, these findings suggest that digital skills are not uniformly prone to decay and therefore training can be made more effective and efficient by targeting the most perishable skills. Different methods for improving skill retention are discussed.

ABOUT THE AUTHORS

Dr. Gregory Goodwin is a Research Psychologist at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI). He holds a Ph.D. in Psychology from Binghamton University.

Dr. Jean Dyer is a Research Psychologist and Team Leader at ARI. She holds a Ph.D. in Research Design and Statistics from Michigan State University.

Dr. Bruce Leibrecht is a senior research psychologist with substantial experience in digital training, digital proficiency assessment, and human dimensions of digitization. He holds a Ph.D. in Experimental Psychology from Michigan State University.

Rich Wampler is a retired Army officer who continues to conduct and support various training research, analysis, and development projects to benefit Soldiers.

The Training and Retention of Selected FBCB2 Operator Skills

Gregory Goodwin, Jean Dyer
U.S. Army Research Institute
Fort Benning, Georgia
gregory.goodwin1@us.army.mil, jean.dyer@us.army.mil

Bruce Leibrecht, Richard Wampler
Northrop Grumman Corp.
Columbus, Georgia
bruce.leibrecht@ngc.com, rich.wampler@ngc.com

INTRODUCTION

The Army Battle Command System (ABCS) is a computerized command and control system. The ABCS family includes the All Source Analysis System (ASAS) used by the intelligence staff; the Advanced Field Artillery Tactical Data System (AFATDS) used by fire support elements; and the Force XXI Battle Command Brigade and Below (FBCB2) or Blue Force Tracker (BFT), the satellite version of FBCB2, used in vehicles and command posts for tracking vehicle locations.

These systems allow leaders to share information including precise positions of friendly and enemy units, graphics, overlays, reports, and orders, over a tactical digital network (Seacord, 2000).

In theory, the ABCS is to be a force multiplier, allowing commanders to react faster, deploy forces more efficiently, and ultimately plan and make better decisions than their adversaries. Unfortunately, the full potential of these systems has not been realized (Clark, 2005). Numerous problems including non-standard hardware, software incompatibilities, and training that cannot keep pace with software upgrades have all limited the effectiveness of these systems (Clark, 2005).

In addition to these challenges, Soldiers and leaders have found that their proficiency at operating these systems is perishable if they do not regularly work with them. Much of the evidence for this comes from anecdotal reports by various leaders (e.g., Lynch, 2001; Johnston, Leibrecht, Holder, Coffey, & Quinkert, 2002) and analysis of training exercises (e.g., U.S. Army Armor Center, 1996).

Reports of the perishability of digital skills agree with what psychologists know about discrete procedural skills (a category into which most digital skills fall). Namely, such skills are easily forgotten relative to continuous skills (Adams, 1987). Discrete procedural skills are differentiated from continuous skills in that the former have a distinct beginning and end. Continuous skills are skills like riding a bike, which, as the proverbial wisdom goes, are rarely forgotten once learned.

In an aircraft cockpit, for example, pilots use checklists to remember discrete tasks like engine startup or takeoff procedures. Without such checklists, research shows these procedures are quickly forgotten, but the ability to maneuver the aircraft, a continuous skill, shows virtually no decay over long periods of time (Schendel, Shields, & Katz, 1978).

Despite this agreement between anecdotal reports of digital skill decay and what psychologists know about discrete procedural skills, there is very little experimental evidence documenting skill decay on ABCS systems. Even worse, the few studies available are equivocal. At least one report documents little skill decay among ASAS operators (Schaab & Moses, 2001) while another indicates more substantial decay (23%-52%) using the Inter-vehicular Information System (IVIS), a vehicle mounted system that predates FBCB2 (Sanders, 1999).

To date, no controlled, experimental data have documented the rate of forgetting of tasks on the current ABCS suite of systems. To begin to fill this gap in our knowledge of the decay rates of digital tasks, we assessed knowledge and operator skills in a group of Army officers who attended FBCB2 familiarization training.

METHOD

Participants

Participants were officers attending the Infantry Captains Career Course (ICCC) at Fort Benning, Georgia during the spring and summer of 2006. All participants had just completed 16 hours of FBCB2 familiarization training. During the two days of this familiarization training, officers received hands-on training of common FBCB2 functions using a desktop computer.

A total of 77 officers participated in the baseline measurement session and 54 (70%) returned for the recall measurement session. For purposes of simplicity in data analysis, only the data from the group that participated in both measurement sessions were analyzed.

An examination of the 23 Captains who were only tested at baseline revealed no distinctive characteristics of this group. Those tested only at baseline were comparable to the rest of the sample in terms of FBCB2 training and experience, self-rated FBCB2 proficiency, general computer experience, and baseline performance on the FBCB2 knowledge and hands-on tests. Comparisons across the two groups with independent sample t-tests and chi-square analyses did not reveal significant differences on any of these measures.

Measurement Instruments

Participants completed three separate instruments in both the baseline and recall measurement sessions: a questionnaire regarding experience and training on digital systems, a test of knowledge of various FBCB2 functions and capabilities, and a hands-on ability test of 13 FBCB2 tasks.

Experience and training questionnaire. For the baseline session, participants were asked a set of five questions regarding their experience and training on ABCS systems. In the first question, participants indicated types of individual operator training they had received and the hours of instruction for each type of training. Types of training included online courses and new equipment training. In the second question, they completed a checklist to indicate the types of collective training they had received on ABCS systems. Types of collective training included motor pool training, and various field training or command post exercises. In the third question, participants listed the systems they had used while deployed to a combat theater, their duty positions while using the system and the number of months they used the system.

In question four, participants rated their overall proficiency on ABCS systems on a four-point scale. The levels of the scale were: 0 - never used, 1 - basic, 2- medium, 3 -high ability. At a basic level, individuals were saying that they could use the system to perform a limited set of functions. At the medium level participants were saying that they were knowledgeable about most of its functions and had limited troubleshooting experience. At the high level, participants were saying that they had advanced knowledge of the system and were often asked for help by others.

In the final question, participants completed a checklist indicating general computer experience on Windows, Macintosh, and Linux operating systems. To reduce the responses in this checklist to a single number, the total number of checks across all operating systems was tallied for each participant. This score, composite computer experience, could range from zero (indicating no experience on any of the three systems) to twenty-four (indicating extensive experience using all operating systems).

In the recall session, the training and experience questions were not repeated. None of the participants indicated that they had used FBCB2 during the period between the two testing sessions.

Knowledge test. The knowledge test consisted of nine questions. The test was designed by a subject matter expert (SME) and validated by the FBCB2 instructors. The items on this test were identical for the baseline and recall tests. participants were asked to name or explain various aspects of the FBCB2 system (e.g., name the four main areas of the operations screen.)

All of the questions were fill-in-the-blank except one. Six of the questions had multiple parts (e.g., “Name the 2 screens on the FBCB2 system.”). In such cases, participants were given partial credit for each component that was answered correctly. All knowledge tests were then scored independently by two raters and the responses were compared. Differences were resolved by discussion among the two raters. The nine questions involved a total of 22 separate responses on the knowledge test.

Hands-on test. The participants completed a 13-item, hands-on test in both the baseline and recall sessions. Each participant used a PC running the FBCB2 software to complete the tasks. All participants completed the hands-on test individually. There was no time limit.

As with the knowledge test items, the items on the hands-on test were initially developed by an SME with experience teaching FBCB2 to Soldiers. The FBCB2 instructors at Fort Benning verified that test items were covered during the two-day training received by the ICCC students.

To score the hands-on test, a data collection sheet was developed. This form contained a series of objective questions for each task on the test. Scoring was based on observable files, folders and settings created by the participant taking the test. For example, in task 3, participants were asked to position their icon on the map. On the data collection sheet, an experimenter indicated two things: a) whether the icon was present on the map, and b) whether the grid coordinates of the icon's location were correct. Thus, this item had a maximum score of two. Whenever an error occurred, a description of the error was recorded on the data collection sheet. All computers were independently checked by two data collectors. When there was a disagreement, both raters looked at the system and came to a consensus.

The baseline and recall tests had the same tasks and task order, but some details of the tasks were altered on the recall test to avoid potential recall of outcomes from memory (e.g., coordinates, file names and settings were varied). Only the wording for three tasks remained completely unchanged across the two tests.

During the baseline session, several participants were observed using the online FBCB2 help function. We did not anticipate this possibility on the baseline measurement so an item was added to the recall test asking participants to indicate whether or not they used the online help function for each task.

Procedure

The baseline test occurred at the conclusion of two days of FBCB2 familiarization training, and the second measurement followed exactly eight weeks later. The two-day training was part of the normal ICCC program of instruction. During the FBCB2 training, the instructors covered start-up and shut-down procedures, and then all of the major functions of the system. Typically, an instructor would demonstrate a procedure while the students repeated the steps on their own system, then the instructor would have students complete a practical exercise on their own.

Testing took place in digital classrooms. All classrooms were identical and had about 44 computer workstations with FBCB2 software installed. The monitors were mounted below a plate glass desktop and faced upward so they could be easily viewed. This configuration also made it difficult for participants to see any adjacent monitors.

Participants completed the experience and training questionnaire and the three measurement instruments: the experience and training questionnaire, knowledge test, and hands-on test. Before completing the hands-on test, participants were given additional instructions explaining how to record their own start and stop times for the hands-on tasks using a digital clock displayed on an overhead projector. It was necessary to have them record their own start and stop times because all of the participants executed the tasks at their own pace. A demonstration by the experimenter helped to illustrate this procedure.

Participants were told to proceed at their own pace and to raise their hands when they completed all 13 tasks. As they completed the hands-on test, an experimenter then verified that they were properly answered. The recall session was conducted exactly like the baseline session. After all officers were released, two experimenters checked each system. Discrepancies were resolved before powering down any system.

RESULTS

Experience and Training

Individual training occurred more frequently on FBCB2 than the other systems, although only a minority of participants (less than 20%) received any given type of training (see Table 1). The greater percent receiving training on FBCB2 probably reflects the fact that students were infantry officers who had led and commanded at the platoon and company levels where FBCB2 is predominantly used. The most frequent type of individual training for FBCB2 was "other" which was consistently defined by participants as on-the-job training received while they were deployed to Afghanistan or Iraq.

A minority of the participants (30% or less) received collective unit training on a specific ABCS system. As with individual training, FBCB2 was the system for which most individuals received any collective training. Motor pool training with FBCB2, the most frequently reported type of collective training, was claimed by only 30% of the participants. Field training exercises (FTXs) and CPXs accounted for most of the remaining collective unit training received by the participants on FBCB2.

Nearly three quarters of the participants (72%) used FBCB2 in a combat theater, while only a small percent (6% or less) used any of the other systems in combat Most veterans gained their experience in Iraq (81%) rather than Afghanistan (19%).

Table 1
Percent (%) of Participants Who Received Individual Training on Digital Systems

	System				
Training	None	FBCB2	ASAS	MCS	AFATDS
Online Course	100	0	0	0	0
Officer Basic Course	90	4	0	0	6
NET	87	11	0	2	0
NET Delta	98	2	0	0	0
Digital Master Trainer	100	0	0	0	0
Other	82	17	0	2	0

Note. NET = New Equipment Training, NET Delta = NET on software changes only. "Other" training occurred on-the-job while deployed to Iraq or Afghanistan.

Consistent with their training and experience, which favored the FBCB2 system, 89% of the participants rated themselves at either a basic or medium level of proficiency on FBCB2. In contrast, fewer than 15% rated themselves at a basic or medium level of proficiency on any of the other systems (see Table 2). Nobody rated themselves at a high level of proficiency.

Table 2
Self-Ratings of Operator Ability (% of Participants)

Self-rating	FBCB2	ASAS	MCS	AFATDS
never used	11	98	87	93
basic	59	2	9	7
medium	30	0	4	0

All participants had used Windows computers while 28% said they had used Macintosh computers and 19% said they had used a Linux computer. Over half of participants had installed software or patches. Almost one third had authored web pages or changed boot-up options and altered BIOS settings.

Knowledge Test Performance and Retention

Overall performance analysis. The responses to the knowledge test were analyzed from the standpoint of (a) the total number of questions answered correctly (i.e., all components of a question had to be answered correctly to get credit for the question) and (b) the total number of components answered correctly. There were nine questions so the question score could range from 0-9; and there were 22 components so the component score could range from 0-22.

The average question score on the baseline test, was 3.6 (40% correct) and the average on the retention test was 3.4 (38% correct). The component scores were 10.17 points (46.2%) on the baseline test and 10.24 (46.5%) on the recall test. Neither change was statistically significant.

Item analysis. Despite the negligible forgetting overall, there were some sizable changes in performance on individual questions. Table 3 presents the percent of the sample that correctly answered each question on the knowledge test .

Chi-square tests were performed to compare the proportion that answered each question correctly on the recall test as compared to the baseline test, using McNemar's (1975) method for correlated proportions. There was significant forgetting on only one of the questions (# 5, see Table 3). Interestingly, there was a significant improvement in performance on two of the questions, (#7 and #1, see Table 3).

An examination of the errors made on the recall test is useful for understanding what was typically forgotten. For example, in questions 1 and 2, participants were asked to spell out two acronyms (FBCB2 and FIPR [flash, immediate, priority, routine]). Surprisingly, the acronym did not always serve as a mnemonic device. Recall was reasonably good for FBCB2. Those who made mistakes often only missed one word and the mistakes were not too different from the correct word (e.g., battlefield in lieu of battle, or communications in lieu of command). In contrast, when spelling out FIPR, 41% received no credit.

Participants appeared to have difficulty when questioned about facts and figures related to the FBCB2 system. For example in questions 3 and 4, the participants were asked to name FBCB2 screens and areas of the operations screen for which they generally had poor recall (see Table 3). The most common mistakes involved assigning names related to functions like "map", or "maneuver" for the operations screen or "message line" or "message bar" in stead of FIPR queue.

Table 3
Percent (%) of the Sample Answering Each Question Correctly on the Knowledge Test

Question	Baseline	Recall	Difference
5. Packet mode message size limit?	69	24**	-45
3. The two system screens are?	15	4	-11
2. FIPR stands for what?	46	39	-7
10. Time zone to enter reminders?	74	70	-4
9. Advantage of FBCB2-T & BFT?	0	0	0
4. Four main areas of Ops screen?	0	2	2
8. How to speed up a slow system?	67	76	9
7. LOS tool determines what?	48	67*	19
1. FBCB2 stands for what?	43	63*	20
Overall	40	38	-2

Note: FIPR = flash, immediate, priority, routine, FBCB2-T = FBCB2 terrestrial, BFT = blue force tracker, LOS = line of sight.
** $p < .01$, indicating that more individuals decreased than increased. * $p < .05$, indicating that more individuals increased than decreased.

Other questions that were recalled poorly were questions 5 and 9. In question 5, participants performed at chance level during recall when asked to identify the packet mode message size limit (24% correct on a four-choice question) indicating they were guessing. In question 9, they were able to recall general benefits of either system but not advantages specific to one or the other.

On the other hand, participants performed fairly well answering questions that related to their general military knowledge such as question 7 (about the line of sight tool) and question 10 (the time zone to enter periodic reminders). They also performed well on question 8 (how to speed up a slow system), but this may have been because this point was emphasized by the course instructors.

Hands-on Test Performance and Retention

Overall performance analysis. The hands-on test entailed 13 tasks performed on an FBCB2 workstation. Scoring of each task depended on the observable footprint left on the FBCB2 workstation (e.g., checking to see that an icon was in the right location). As with the knowledge test, a task score (analogous to the knowledge test question score) reflected the number of tasks (range: 0 - 13) for which all components were performed correctly. In addition, a hands-on component score (analogous to the knowledge test component score) was derived to reflect the number of task components (range: 0 - 37) completed correctly.

Performance summed across the 13 tasks showed modest (10%) decay. In the baseline, the participants completed an average of 72% of the tasks correctly without errors in any of the components and this dropped to 62% in the recall test. Analysis of variance revealed the test-retest effect was significant, $F(1,53) = 17.07$, $p<.001$. When looking at component scores, less decay was seen. Component scores averaged 78% correct in the baseline test and 73% correct in the retest. This modest (5%) decrease was statistically significant, $F(1,53) = 4.09$, $p<.05$.

Item analysis. Table 4 shows the percent of the sample correctly performing each of the 13 tasks. To compare performance across the two testing periods, McNemar Chi-square analyses were performed. Results indicated that there was a significant decline in performance on three of the thirteen tasks. There was no significant improvement in performance on any of the tasks.

It is important to note that task 12, center the icon on the map, could not be executed unless part of task 3, position the icon on the map, was done successfully. . When only those people who succeeded in placing their icon on the map (whether or not it was in the correct location) for task 12 of the recall test were examined, there was only a small, non-significant decline in their ability to also center the icon (90% on the baseline test vs. 87% on the recall test). Thus, the significant decline in performance on task 12 was primarily due to forgetting how to get the icon on the map, not how center the icon (note: to successfully complete task 3, participants had to get the icon on the map at the specified location but they could center the icon on the map regardless of where it was located.)

Although nearly everyone performed the first task, verify platform role at recall, only 9 could perform task 2, clear logs and queues. It is interesting to note that more individuals used help on this task than on any other task.

Only 34 (63%) of the participants were able to place their icon on the correct map location for task 3. Of those who did not get credit on this task, 14 were not able to place their icon on the map at all and the rest put it in the wrong location.

Most participants were able to perform tasks 4 and 5. When asked to change the situation awareness settings in task four, nine individuals made no changes and three individuals changed the settings, demonstrating they knew what to do, but they made errors in their changes. Only three participants could not create any folders in task 5, and two created only one folder demonstrating they knew the process.

In tasks 6, 7, 8, and 9, participants had to create and/or address various types of messages. An error common to all four of these tasks stemmed from using the search function to find an addressee. The address book search function in FBCB2 does not always return an exact match to the search string when one exists. For this reason, it is always important to double check the search result before selecting the addressee. On all of these tasks, an average of 48% of those who made errors, misaddressed their messages because they did not carefully check the search result before selecting the addressee. Other participants were not able to assign an address at all, or wound up self-addressing the message.

In tasks 7 and 9, participants had to alter the precedence settings of the outgoing messages. In task 7, half, and in task 9 over two-thirds, of those who made errors, failed to set the precedence of the message correctly. Most commonly, participants left the precedence setting at its default value.

For task 10, display and save a message, only eight individuals were unable to save the message and four saved it to the wrong location. For task 11, create a route, only one in four completed it successfully and the most common error (made by 32 individuals) was that participants failed to leave the route displayed. Another common error for task 11 (made by 21 individuals) was to add too many waypoints to the route. Most had only one extra waypoint, probably because they didn't realize their current position was a waypoint.

Table 4

Percent (%) of the Sample Correctly Performing the 13 Hands-on Tasks

Task	Baseline	Recall	Difference
12. Position and Center icon on map[a]	87	63 **	-24
11. Create route on map	46	26 *	-20
6. Create address group	74	55 *	-19
2. Clear logs and queues	33	18	-15
7. Set free text defaults	65	52	-13
3. Position icon on map	76	63	-13
10. Save incoming message	91	78	-13
4. Adjust SA settings	89	78	-11
5. Create message folders	100	91	-9
1. Verify platform role	93	91	-2
8. Create/send free text message	85	85	0
9. Create/send SPOT report	22	24	2
13. Check line of sight	69	72	3

[a] Task 12 was dependent on task 3 because to center the icon on the map, it must first be placed on the map.

* $p < .05$, ** $p < .01$.

Time to Complete Hands-on Tasks

In the baseline test, the participants took an average of 28.0 min total time to perform the hands-on test. The time included the time spent reading directions and executing the tasks. The participants expended an average of 35.1 min in performing the recall test, for a mean increase of 7.1 min. The increase in time was significant, $t(50) = 6.47$, $p < .001$.

In the analysis of the time to complete individual tasks, nine of the thirteen tasks took significantly longer to execute and one took significantly less time (line of sight tool). Thus, time to complete tasks was a more sensitive measure of forgetting than overall success rates.

Use of FBCB2's Help Function

Only during the recall test were the participants asked to record whether they used the help feature for a given task. In that session, 27 participants (50% of the sample) reported using help on at least one task. Half of these indicated they used help two to five times. Those who used help at least once averaged 66.4% correct in the recall test, while those who did not performed significantly better, averaging 78.9% correct, $t(52) = 2.13$, $p < .05$. Those who used help took more time (37.4 min vs. 33.1 min) to complete tasks, though this was not a statistically significant difference.

In only three tasks did more than 10% of the 54 participants report using help: (a) clear logs and queues, (b) position icon on map, and (c) create an address group. In the case of clearing logs and queues, 31% of the participants reported using help. Interestingly, only 18% of the sample performed this task correctly on the recall test, the lowest percentage of all 13 tasks.

Self-Rated Proficiency Groupings

Performance on the knowledge and hands-on test varied as a function of self-rated proficiency. Those who had never used FBCB2 before performed significantly worse than participants who rated themselves at either a basic or medium level of proficiency. (knowledge question score, $F[2,51] = 4.8$, $p = .01$; knowledge component score, $F[2,51] = 5.1$, $p = .01$; hands-on task score $F[2,51] = 3.7$, $p = .03$; and hands-on component score, $F[2,51] = 4.0$, $p = .03$). Those who had never used FBCB2 also showed more skill decay on the hands on test ($F[2,51] = 3.6$, $p = .04$, component score).

Experience and Training Predictors of Performance

Total hours of FBCB2 training, total collective training experiences, total months of FBCB2 use in combat, self-rated FBCB2 proficiency, composite computer experience, and baseline knowledge test score were all entered into a multiple regression equation to predict performance on both component and task scores for the hands-on tests. The component recall score ($R = .56$, $p < .01$), and the baseline ($R = .50$, $p < .05$) and recall ($r = .57$, $p < .01$) task scores were all significantly predicted with these measures. This means that collectively the experience and training measures accounted for between 25% and 32% of the hands-on performance score.

DISCUSSION

Retention of Digital Operating Skills

Over the eight-week retention interval, performance declined significantly on 3 of the 13 hands-on tasks. Although there was no significant change in the overall scores on the knowledge test, the percent of the sample giving correct answers declined significantly on one of the questions and improved significantly on two others. This unexpected increase in knowledge is most likely due to coincidental learning that took place during the retention interval (i.e., from classroom activities or discussions with colleagues).

The overall decline in proficiency on the hands-on test in the present research was 10% for the task scores and 6% for the component scores (both declines were statistically significant). By comparison, in the investigation of IVIS skills by Sanders (1999), decay over a 30-day no-training period was considerably larger (23% for message skills and 52% for overlay skills). Although future research will be needed to definitively explain these discrepant results, there were differences between the participants and methodologies that might explain them.

One possible explanation has to do with differences in the system-specific experience of the participants in two experiments. Almost 72% of the sample in the present experiment reported using FBCB2 during deployment in Iraq or Afghanistan. By comparison, none of Sanders' participants reported previous training on IVIS. This stronger baseline knowledge of FBCB2 may have made the participants' skill in the present study more resistant to decay. Recall that participants who indicated that they had no prior experience using FBCB2 did worse on the hands-on test and forgot significantly more than those who rated themselves at a basic or medium level of proficiency.

In a related vein, participants in the current study indicated substantial general computer experience. Nearly all individuals (94%) reported using software applications in a Windows environment, and most (81%) had installed application software. In contrast, fewer than half of Sanders' participants (39%) had used a computer for more than a year, and one in four said they did not use a computer at all.

Finally, the fact that Sanders used only individuals who performed to criterion (i.e. demonstrated that they were proficient on all tasks tested) is another likely reason that he reported greater skill decay. Because his participants started at 100% proficiency, they had a greater potential to show decline. Participants in the present study often started at low levels of proficiency and had less potential to show decline because of a floor effect.

Understanding Performance Errors

On the knowledge test, memorizing facts with low meaning or connectedness made for a difficult learning challenge. A good example of hard-to-memorize information was found in the names of the four main areas of the FBCB2's Operations screen. These names (classification banner, map area, operations function bar, and FIPR bar) have low intrinsic meaning and do not relate to terminology from a common system such as Microsoft Windows. Recall of these names reached only 22% at the end of the two-day course. Another example was the size limit for messages sent in packet mode, which exhibited 45% decay. These findings fit with the literature on verbal learning, where low meaningfulness of the subject material is well known to impede learning (Underwood, 1966).

The number of elements in a question did not relate significantly to performance on the knowledge test. The same was true for the number of steps in a hands-on task. This was somewhat surprising, in light of the literature relating task complexity to skill retention (Goodwin, 2006). Although this may mean that the number of steps is not a good indicator of task complexity, it is also possible that because task complexity diminished both initial performance and recall performance, a floor effect made it difficult to detect decay. For example, the task of create and send a SPOT report, which had more steps than any other, was performed by only 22% of the sample at baseline and 24% at recall. With such a low level of performance at baseline, it is difficult to detect decay. Clearly it would be better to have everyone start with proficiency on all tasks at baseline.

Looking at the hands-on task components that had greater than average (11%) change, highlights some of the more notable recall problems. For example, placing the icon on the map and doing so at a specific coordinate showed 28% and 13% decay respectively. Ordinarily with a functioning GPS, such a task would be unnecessary as the system would automatically position the vehicle icon on the map at the correct coordinate. It is possible that because there was seldom a need to perform this function in theater, it was more easily forgotten. Another factor contributing to decay may be that the first step to placing your own vehicle icon on the map is to press the F6 Admin button, which is less intuitive than, for example, the F1 Map button might be.

Another map component task that showed greater than 11% decay was leaving the route displayed on the map. The steps required to perform this step are not entirely intuitive although the question provided a hint with the phrase "leave the driver's display on." This prompt resembled the checkbox labeled "driver's display on." that needed to be checked to accomplish this task. Additional research is needed to understand why this was forgotten.

The most common messaging problem stemmed from a quirk of the address book search function. The addressee search function does not prioritize whole matches over partial matches. So when the addressee (CDR-3ID-BDE1) was entered, the search function returned the first addressee in the list containing that string (DCDR-3ID-BDE1). In fact, the addressee was chosen to determine whether participants would check the addressee returned by the search function. Failing to check the addressee was a major contributor to errors on two of the component tasks (i.e. creating an address group and setting the free text message defaults). Future versions of the FBCB2 software might consider modifying the search function, but in the mean time, this is a point that needs training emphasis.

Another feature of this system that may have contributed to errors has to do with the message options for the SPOT report. This dialogue box has two tabs, one to set the message precedence and one to set the message addressee. Clicking on either tab brings it to the foreground and settings can be changed for that tab in a way that is analogous to a Windows system. At the bottom of the message options dialogue box are several buttons visible no matter which tab is in the foreground. Those include "okay", "apply", and "close" buttons. In a Windows system, changes can usually be made to any tab in any order and then a single mouse click on "apply" or "okay" will accept those changes, but this is not the way it is done with FBCB2. The "apply" button must be selected after changes are made to each tab. If another tab is selected before selecting "apply", all the changes made to the prior tab will resort to their default settings. It is likely that errors setting the precedence for the free text message and the SPOT report are because participants forgot about this idiosyncrasy.

After sending a SPOT report, clearing logs and queues was performed by the lowest percentage of individuals on the baseline test and had the lowest rate of all tasks on the recall test. This task had only five steps but screen prompts for this task are largely missing. Performing this task required the user to first choose the "start" button in the lower left corner of the screen, then "FBCB2" and then "clear logs and queues", followed by selecting items to clear and then selecting the "apply" button. Forgetting on this task was most likely due to failure on one of the first two steps as it is hard to imagine that someone who completed those two steps would be unable to complete the rest.

Techniques for Counteracting Decay

The outcome of the present study indicates that there are some ways current trainers might improve training and retention for FBCB2 operator skills. First, because we found significant effects of self-rated proficiency on learning and recall, it could be beneficial to identify those individuals who have had no prior exposure to FBCB2 for some remedial training. Those individuals indicating they had never used FBCB2, learned less and forgot it faster than those who gave themselves low or medium self-ratings of proficiency.

Accommodating these differences in a single group of students can be challenging, but doing so is important for getting everyone to a higher level of proficiency (Wampler, Dyer, Livingston, & Blankenbeckler, 2006). Those with no previous exposure to the system may be especially motivated to learn, a factor upon which instructors can capitalize. The digitally savvy students can buddy with less experienced colleagues to serve as demonstrators and discussion group participants. The inexperienced students can be encouraged to seek help from the instructors or peer coaches, both during and after a training session.

It was also clear that some topics and tasks were easily forgotten by the students; therefore, developing mnemonics or other memory aids, or perhaps reallocating time to focus on some of the more easily forgotten topics and tasks could help counteract their tendency to be forgotten. For example, instructors might develop mnemonics for the names of the two main FBCB2 screens, the largest allowable message size, the four main areas of the operations screen, and the advantages of FBCB2 and BFT. Similarly, greater emphasis needs to be given to helping students to remember the steps for clearing logs and queues, creating a SPOT report, selecting addressees using the search function, and manually placing their icon on the map.

Still other training emphasis might be added to the use of the FBCB2 help function. As reported above, use of the help system was associated with lower success on the hands on tasks. Although this most likely indicates that those who used the help function were those who were least knowledgeable about the system, one might expect that the help function would have been more effective at compensating for those deficiencies.

In addition to altering training, future versions of FBCB2 software could be redesigned to help attenuate errors on tasks. For example, the addressee search function might be redesigned to prioritize whole matches over partial matches. Another software redesign should allow changes made to different tabs within a dialogue box to be preserved with a single click of either the "apply" or "okay" button rather than requiring these buttons to be selected before changes can be made to another tab. Finally better on-screen prompts might be added to cue users how to clear logs and queues and to place their icon on the map.

Conclusions

Decay of the FBCB2 skills investigated in this report was observed over an eight week retention interval. The good news is that overall performance only declined by 10%. In fact, there was significant decay on only three of the 13 tasks, each one with 19% - 24% decline. When looking at the components of those tasks, it appears that most of the decay was associated with forgetting how to place an icon on the map, forgetting how to leave a route displayed, forgetting how to assign the correct addressee, and forgetting how to clear logs and queues. Elapsed time to perform hands-on tasks appears to be a more sensitive measure of decay revealing skill decline on eight of the thirteen tasks but improvement for one task.

Understanding the causes of the decay on these tasks is still speculative, but it is unlikely that a single problem accounts for all forgetting. The background experience and training of individuals contributes to decay, predicting about 25% - 32% of performance. Vague or absent screen prompts, system idiosyncrasies, and an inability to use the help system effectively also account for some decay although a precise number can't be attributed to those factors just yet.

Trainers may want to devote additional instructional time to counteract these problems and system designers may want to make changes to the interface. It is important to keep in mind, however, that this is the first empirical analysis of FBCB2 skill decay so future studies will need to confirm these results before a strong recommendation can be made to alter training programs of instruction or the system interface. Future research should also include a wider range of tasks representing current operational uses of the system.

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.

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

Goodwin, G. A. (2006). *The training, retention, and assessment of digital skills: A review and integration of the literature* (ARI Research Report 1864). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Johnston, J. C., Leibrecht, B. C., Holder, L. D., Coffey, R. S., & Quinkert, K. A. (2002). *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. (DTIC No. AD-A412717)

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. (DTIC No. AD-A395042)

McNemar, Q. (1975). *Psychological statistics* (5th Ed.). New York: Wiley.

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. (DTIC No. AD-A368212)

Schaab, B. B., & Moses, F. L. (2001). *Six myths about digital skills training* (ARI Research Report 1774). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD-A392922)

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. DTIC No. AD-A061338)

Seacord, C. R. (2000). *Digitizing the battlefield* (U.S. Army War College). Carlisle Barracks, PA: U.S. Army War College. (DTIC No. AD-A377740)

Underwood, B. J. (1966). *Experimental psychology*. New York: Appleton-Century-Crofts.

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

Wampler, R. L., Dyer, J. L., Livingston, S. C., Blankenbeckler, P. N., Centric, J. H., & Dlubac, M. D. (2006). *Training lessons learned and confirmed from military training research* (ARI Research Report 1850). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD-A446697)