

Technical Report 1160

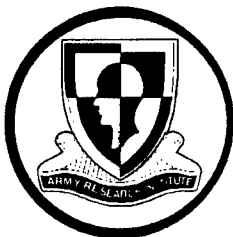
**Optimizing the Speed, Durability, and Transferability
of Training**

**Alice F. Healy, James A. Kole, Erica L. Wohldmann,
Carolyn J. Buck-Gengler, James T. Parker, and
Lyle E. Bourne, Jr.**

University of Colorado at Boulder

April 2005

20050623 121



**United States Army Research Institute
for the Behavioral and Social Sciences**

Approved for public release: distribution is unlimited

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

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

**ZITA M. SIMUTIS
Director**

Research accomplished under contract
for the Department of the Army

University of Colorado at Boulder

Technical Review by

David P. Costanza, The George Washington University
Stephen Goldberg, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Technical 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 Technical 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-YYYY) July 2003		2. REPORT TYPE Final		3. DATES COVERED (From - To) August 1999 – July 2003	
4. TITLE AND SUBTITLE Optimizing the speed, durability, and transferability of training				5a. CONTRACT NUMBER DASW01-99-K-0002	
				5b. PROGRAM ELEMENT NUMBER 611102	
6. AUTHOR(S) Alice F. Healy, James A. Kole, Erica L. Wohldmann, Carolyn J. Buck-Gengler, James T. Parker, and Lyle E. Bourne, Jr. (University of Colorado at Boulder)				5c. PROJECT NUMBER B74F	
				5d. TASK NUMBER 2902	
				5e. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Colorado Campus Box B-19 Boulder, CO 80309				8. PERFORMING ORGANIZATION REPORT NUMBER 153-1392 (35993)	
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'S ACRONYM(S) ARI	
				11. MONITOR'S REPORT NUMBER(S) Technical Report 1160	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES COR: Dr. Paul Gade					
14. ABSTRACT Our research program aims to develop principles that optimize simultaneously all three characteristics of training – speed, durability, and transferability of learned knowledge and skills. Such simultaneous optimization would not necessarily optimize any one characteristic alone but would require instead a balanced consideration of all three characteristics. The balance of the characteristics of training is not fixed across tasks or even within a given task but rather can depend on a variety of external factors, such as fatigue and information load, that can change over time. Two studies in our program are summarized to illustrate our work. The first part of this research involves a data entry task, focusing on initiation and execution of response components under fatigue produced by prolonged work. This research demonstrates that prolonged work affects the component cognitive and motoric processes of data entry differentially and at different points in time. The second part of this research involves a duration estimation task which is in some cases coupled with a secondary articulatory suppression task. It focuses on ways to promote transfer of training. This research demonstrates that learning how to estimate durations is highly specific to the conditions of training and critically depends on whether or not a secondary task is required.					
15. SUBJECT TERMS Training, transfer, long-term retention, knowledge and skills					
SECURITY CLASSIFICATION OF:			19. LIMITATION OF ABSTRACT Unclassified	20. NUMBER OF PAGES 23	21. RESPONSIBLE PERSON Ellen Kinzer Technical Publication Specialist 703-602-8047
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified			

Technical Report 1160

Optimizing the Speed, Durability, and Transferability of Training

**Alice F. Healy, James A. Kole, Erica L. Wohldmann,
Carolyn J. Buck-Gengler, James T. Parker, and Lyle E. Bourne, Jr.**
University of Colorado at Boulder

**Research and Advanced Concepts Office
Paul Gade, Chief**

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

April 2005

**Army Project Number
0601102A**

**Personnel, Performance
and Training**

Approved for public release; distribution unlimited.

FOREWORD

The mission of the U. S. Army Research Institute for the Behavioral and Social Sciences (ARI) is to maximize individual and unit performance and readiness to meet the full range of world-wide Army missions through advances in the behavioral and social sciences. The purpose of this document is to describe the work on one ARI research project that develops principles of training that promote efficient learning, durable memory, and flexible transfer performance through an understanding of cognitive functioning and its role in improving training. Contrary to popular belief, findings from this research effort show that to be effective, training must incorporate as many of the complete set of field task requirements as possible, including all secondary task requirements that might be imposed. Furthermore, this research demonstrates that, although introducing sources of interference into a task or increasing the difficulty of the task slows down initial skill acquisition, these variations ultimately lead to improvements in the durability and flexibility of the learned skill. The findings from this basic research effort are already being transitioned for operational training assessment in the simulator systems applied research program.

A handwritten signature in black ink, reading "Michelle Sams". The signature is fluid and cursive, with the first name "Michelle" and the last name "Sams" clearly distinguishable.

MICHELLE SAMS
Technical Director

Acknowledgment

This document serves as a final report for Army Research Institute Contract DASW01-99-K-0002 (Alice F. Healy, Principal Investigator; Lyle E. Bourne, Jr., Co-Principal Investigator). The research reported here was also supported in part by Army Research Institute Contracts DASW01-96-K-0010 and DASW01-03-K-0002 (Alice F. Healy, Principal Investigator; Lyle E. Bourne, Jr., Co-Principal Investigator), and supplemental support was received from Army Research Office Grant DAAG55-98-1-0214 to the University of Colorado (Alice F. Healy, Principal Investigator).

OPTIMIZING THE SPEED, DURABILITY, AND TRANSFERABILITY OF TRAINING

EXECUTIVE SUMMARY

Research Requirement:

The U. S. Army spends much time and many resources in training its personnel. Training is essential because recruits cannot be expected to come equipped with the military knowledge and skills they will need in the field. But training is costly, so it is important to insure that it be accomplished as quickly and as efficiently as possible. Increasing training speed, however, should not be the only, or even the most important, consideration. If soldiers have successfully learned how to perform a task during training but then forget how to perform it at the time that they need to do so, the training has clearly been inadequate. Passing a test at the end of training does not guarantee later success in the field. Training needs to be durable as well as efficient. But even durable training cannot guarantee that learned knowledge and skills will be applied successfully to situations different from those encountered during training. Training circumstances can rarely capture the full set of circumstances under which tasks are encountered in the field. It is, therefore, essential that training be transferable as well as durable. It is the aim of our research program to develop principles that separately optimize the three major aspects of training: (a) its speed or efficiency, (b) its durability or long-term retention, and (c) its transferability or generalizability to new situations.

Procedure:

Although many of our studies have overlapping goals, in the present report we focus on a subset of our studies and discuss two separate topics. The first topic involves managing factual overload, rapidly presented information, stress, frustration, and fatigue, with an emphasis on tasks involving perceptual and motoric processing. The experiments reported for this topic focus on the specific issue of initiating and executing response components under fatigue produced by prolonged work. The second topic involves optimizing the balance of the three major aspects of training. The research reported for this topic focuses on the specific issue of ways to promote transfer of training.

Findings:

Managing factual overload, rapidly presented information, stress, frustration, and fatigue. When participants work continuously over time on a task, such as data entry, two opposing processes (facilitative and inhibitory) might affect their performance. On the one hand, performance might improve, becoming more accurate, faster, or both, as participants master the skills required of data entry. On the other hand, either or both aspects of performance might deteriorate as participants suffer the effects of fatigue and boredom over long trial periods. We completed two experiments addressing fatigue in a repetitive data entry task. Under conditions promoting fatigue, participants entered four-digit numbers on a computer terminal. In Experiment 1, accuracy worsened but response times improved both across and within session halves, reflecting an increasing response latency-accuracy tradeoff. In Experiment 2, the (largely

cognitive) time to enter the first digit of each number improved over the first session half but worsened over the second half. Accuracy worsened but time to enter the remaining digits improved across though not within session halves. The (purely motoric) time to press the enter key improved across and within session halves. Thus, through a combination of practice and fatigue, prolonged work affects the component cognitive and motoric processes of data entry differentially and at different points in practice.

An additional experiment investigated effects of articulatory processing on number data entry. Participants entered four-digit numbers presented as either words or numerals on a keyboard either under articulatory suppression or in silence. The articulatory suppression group typed initial digits *faster* than the silent group, but for subsequent digits, the opposite pattern occurred at least with word stimuli. Thus, articulation of numbers, which promotes entry into the phonological loop of working memory, retards processing of initial digits but enhances processing of subsequent digits.

Optimizing the balance of training aspects. Experiments examined training, retention, and transfer of a duration estimation skill using an arbitrary unit of time in a prospective, production estimation paradigm. Participants were trained with feedback and then either tested immediately for transfer without feedback or retrained with feedback 1 week later. There were three training and retraining conditions, two involving secondary tasks. Retention of the estimation skill was perfect across the 1-week delay when the secondary task condition was unchanged, but there was no transfer of the skill when the condition was changed. These findings are interpreted within the procedural reinstatement framework with the assumption that the duration estimation procedures incorporate requirements of the secondary task.

Utilization of Findings:

Contributions to basic science. The many research directions we have pursued address a diverse set of issues but have two primary goals: First, we intensively examined extraneous variables (including factual overload, rapid presentation of information, stress, frustration, and fatigue) that might adversely affect training, in order to develop procedures to counteract their deleterious effects. Second, building on our previous work which revealed that long-lasting knowledge and skills are highly specific to the training conditions, we conducted a series of studies concerned with optimizing the relationship between all three major aspects of training – its speed, durability, and transferability. These studies shared a common analytical and experimental methodological approach and the common theoretical goal of understanding the psychological principles underlying the acquisition, retention, and generalization of knowledge and skills.

Potential applications. These studies also shared the common applied goal of improving the training of military personnel such that the knowledge and skills will be acquired quickly and yet still be accessible across long delay intervals with no practice and adaptable to new situations outside the training environment. The balance of the three aspects of training (speed, durability, and transferability) is not necessarily fixed across tasks or even within a given task but rather may depend on a variety of external factors, such as stress, frustration, fatigue, rate of information presentation, and information load, that can change from time to time. Variations in

any one of these factors can affect the interaction of these aspects of training. Hence, our studies also examined the three aspects of training under various conditions of stress, frustration, fatigue, rate of information presentation, and information load. This examination should help with designing training programs that will allow the trainee to contend with these external conditions that obtain unpredictably in the field and could adversely influence performance if not properly managed.

OPTIMIZING THE SPEED, DURABILITY, AND TRANSFERABILITY OF TRAINING

CONTENTS

	Page
INTRODUCTION	1
Initiating and Executing Response Components Under Fatigue Produced by Prolonged Work	1
Ways to Promote Transfer of Training	7
Conclusions.....	11
REFERENCES	13

LIST OF TABLES

TABLE 1 ACCURACY (PROPORTION CORRECT) AND RESPONSE TIME (RT) IN MS (TOTAL RT, INITIATION TIME, EXECUTION TIME, AND CONCLUSION TIME) DURING TRAINING IN EXPERIMENTS 1 AND 2 OF HEALY, KOLE, BUCK-GENGLER, AND BOURNE (2004) BY SESSION HALF AND BLOCK	3
TABLE 2 TOTAL RESPONSE TIME IN MS IN EXPERIMENT BY KOLE, HEALY, AND BUCK-GENGLER (IN PRESS) DURING TRAINING BY GROUP, PRESENTATION FORMAT, AND BLOCK.....	6
TABLE 3 MEAN PROPORTIONAL ABSOLUTE ERROR IN EXPERIMENT BY HEALY, WOHLDMANN, PARKER, AND BOURNE (2002) DURING CRITICAL BLOCKS	10

Introduction

Many educational, military, and industrial organizations spend much time and many resources in training their personnel. Because training is costly, it is important to insure that it be accomplished as quickly and as efficiently as possible. However, increasing training speed should not be the only consideration. If individuals have successfully learned how to perform a task during training but then forget how to perform it at the time that they need to do so, the training clearly has been inadequate. Training also needs to be durable. But even durable training cannot guarantee that learned knowledge and skills will be applied successfully to situations different from those encountered during training. It is, therefore, essential that training also be transferable. It is the aim of our research program to develop principles that optimize separately and in combination the three major aspects of training: (a) its speed or efficiency, (b) its durability or long-term retention, and (c) its transferability or generalizability to new situations.

In some of our earlier studies, we discovered that training that minimizes the time to acquire knowledge or skills may be detrimental to long-term retention (e.g., Schneider, Healy, & Bourne, 1998, 2002; Schneider, Healy, Ericsson, & Bourne, 1995). In other words, what is learned quickly often is not what is remembered best. Likewise, in other studies, we found that training that maximizes long-term retention of material may severely limit the transferability of that material (e.g. Healy & Bourne, 1995). For example, being able to retain a fact does not guarantee that its relevance will be recognized in new situations. Thus, in designing an effective instructional program, the goal should be to optimize *simultaneously* all three characteristics of eventual performance – the speed, durability, and transferability of training – taking into account the tradeoffs among them. Such simultaneous optimization would not necessarily optimize any one characteristic individually but would require instead a balanced consideration of all three characteristics.

The balance of the three aspects of training (speed, durability, and transferability) is not necessarily fixed across tasks or even within a given task but rather may depend on a variety of external factors, such as stress, frustration, fatigue, rate of presentation of information, and information load, that can change from time to time. Variations in any one of these factors can affect the interaction of these aspects of training.

The studies summarized in this report illustrate our recent work on two topics. The first topic involves managing factual overload, rapidly presented information, stress, frustration, and fatigue, with an emphasis on tasks involving perceptual and motoric processing. The experiments reported for this topic focus on the specific issue of initiating and executing response components under fatigue produced by prolonged work. The second topic involves optimizing the balance of the three major aspects of training. The research reported for this topic focuses on the specific issue of ways to promote transfer of training.

Initiating and Executing Response Components Under Fatigue Produced by Prolonged Work

We have been exploring the underlying causes of durability and specificity of skill using the data entry task, in which participants see numbers and type them into a computer. Our earlier work on training data entry skills (e.g., Fendrich, Healy, & Bourne, 1991) showed that individuals type numbers that have been typed previously during training significantly faster than numbers that, in the context of the laboratory experiment, are “new” or unfamiliar. This advantage for typing “old” numbers, known as the repetition priming effect, is based largely on implicit or nondeliberate memory. We found that both perceptual and motoric processes contribute to the repetition effect. In more recent work (Buck-Gengler & Healy, 2001), we

demonstrated that the abstract numerical concept, rather than the surface percept, contributes to repetition priming. To manipulate the surface percept the presentation format of the numbers was varied. Half of the numbers were presented symbolically as words (e.g., "two four one seven"), whereas the other half were presented as numerals corresponding to the labels on the data entry keys (e.g., "2 4 1 7").

Participants were trained in the data entry task over a single session and then tested in a second session after a 1-week delay. In both sessions, four-digit numbers were presented one at a time on a computer display screen until the participant responded by typing the number onto the computer keyboard followed by the "enter" key. The presentation and subsequent entry of a four-digit number constituted one trial. Total response times for entering all four digits in a number and the final enter key as well as the accuracy of the entries were recorded on every trial. During the training session, a set of 64 four-digit numbers was used as the learning set. This set was repeated for entry by the participant five times across five blocks of 64 trials during the training session, with a different order of numbers for each block.

The procedure in the testing session was the same as in the training session except that there were only two blocks of trials. Sixty-four "new" four-digit numbers not previously entered by the participants were intermixed with the 64 "old" numbers of the learning set. For half of the old numbers, the presentation format was the same in the testing session as in the training session, and for the remaining half the presentation format in training and testing were different.

One group of participants was trained with the digit keypad to the right of the letter keys on the standard keyboard and was tested with the digit row above the letters on the keyboard. A second group of participants was trained with the digit row and was tested with the keypad. Thus, for both groups, the key configuration used in training did not match the key configuration used in testing to ensure that any repetition effect could not be attributed to the retention of the motoric component of the task.

Numbers presented as numerals were typed more quickly than those presented as words. Also, old numbers were typed significantly faster than new numbers, reflecting repetition priming. There was no overall difference for old numbers between those in the same format and those in different formats at test and at training. However, for both test presentation formats, numbers presented as words in training were faster at test than those presented as numerals, suggesting that having to encode numbers from words in training led to more processing than encoding them from numerals.

In two new experiments (Healy, Buck-Gengler, Kole, & Bourne, 2001; Healy, Kole, Buck-Gengler, & Bourne, 2004), we used a variation of this method to assess the effects of fatigue on data entry performance, to see whether any negative effects of fatigue could be reduced, and to evaluate whether fatigue affects the magnitude of repetition priming.

When individuals work at a continuous task, such as data entry, two opposing processes might affect accuracy and response time. First, the individuals might improve, becoming more accurate, faster, or both in their responses as they master the required skills. Second, one or both aspects of performance might deteriorate as the individuals suffer the effects of fatigue and boredom over long trial periods.

To examine these effects, in Experiment 1 of our fatigue research, there were two sessions in which 32 participants were given strings of four-digit numbers always viewed as numerals displayed on a screen, and they always entered the numbers by typing them on the keypad at the right of the keyboard. In the first session, which was twice the length of that used in the Buck-Gengler and Healy (2001) study, participants saw 64 numbers repeated in different random

orders five times over five blocks in the first half of the session, and they saw five blocks of 64 different numbers in the second half. In the second session 1 week later, there were four blocks of 64 numbers; in each block half of the numbers were repeated from each half of the first session, and half were new. Only right-handed individuals participated, and to promote fatigue, they were required to use their left hand. Also to promote fatigue, the 1/2-s intertrial delay between response and presentation of the next stimulus was eliminated. A short break, up to 5 min in length, halfway through the training session (after the fifth block) allowed participants to have some level of recovery from fatigue.

We found that accuracy significantly decreased from the first to the second half of the session and across blocks within each session half (see Table 1). This finding documents the fact that we successfully induced fatigue during training. Nevertheless, we also found that total correct response time significantly decreased (i.e., improved) from the first to the second half of the session and across blocks within each session half, suggesting that fatigue and practice combined to lead to an increase in the response latency-accuracy tradeoff during training.

Table 1

Accuracy (Proportion Correct) and Response Time (RT) in ms (Total RT, Initiation Time, Execution Time, and Conclusion Time) During Training in Experiments 1 and 2 of Healy, Kole, Buck-Gengler, and Bourne (2004) by Session Half and Block

Measure	First Half					Second Half				
	1	2	3	4	5	1	2	3	4	5
Experiment 1										
Accuracy	.920	.895	.893	.880	.878	.881	.872	.865	.857	.857
Total RT	2718	2699	2688	2657	2635	2644	2607	2543	2573	2537
Experiment 2										
Accuracy	.889	.896	.875	.889	.873	.880	.875	.870	.876	.863
Total RT	2498	2500	2527	2454	2445	2372	2340	2378	2354	2384
Initiation time	1108	1079	1098	1069	1085	1058	1068	1095	1079	1124
Execution time	326	336	333	326	324	306	301	302	303	297
Conclusion time	282	282	284	280	275	265	257	254	253	249

Note. Execution time is the average time per keystroke for the second, third, and fourth digits. Means of medians are provided; thus Total RT is not necessarily equal to the sum of the component RTs.

This experiment yielded significant repetition priming at test 1 week later (i.e., an advantage for old numbers previously entered relative to new numbers) for the measure of total correct response time. There was also a significant interaction of training half and test half on accuracy at test for the old numbers previously entered during training. Accuracy was lowest for those old numbers occurring in both the second half of training and the second half of the test. This finding is consistent with the simple hypothesis that accuracy deteriorates with fatigue so should be worst when the participants' state of fatigue is high during both training and testing.

At the beginning of this research, two opposing ways in which performance might be affected by deliberate practice were proposed. One possibility was that performance should

improve, becoming more accurate, faster, or both as participants learned the task. Alternatively, performance might deteriorate, becoming less accurate, slower, or both as participants suffered from the effects of fatigue. We found that practice showed opposite effects on response latency and accuracy, reflecting an increasing latency-accuracy tradeoff. With practice, responses became increasingly faster but also increasingly less accurate.

Experiment 2 of our fatigue research focused on these findings concerning the increasing latency-accuracy tradeoff with training. The improvement in response times across blocks during training in Experiment 1 could be due to either general or specific training factors. In terms of general factors, the improvement could be due simply to practice with the task. In terms of specific factors, the improvement could be due to within-session repetition priming because each number was repeated five times, once in each of the five blocks within a given half of the training session. To isolate the effects of practice and fatigue and to eliminate the effects of repetition priming within training, in Experiment 2, each number occurred only once, with no numbers repeated during training. Also, to determine whether the decrease in response time and the decline in accuracy across session halves were due in part to peripheral motoric factors involving the specific hands, half of the participants (all of whom were right-handed) switched the hand they used to type from the first to the second half of the session. This *switch* condition was compared to a *no switch* condition in which the same hand was used for typing in both halves. We counterbalanced the hands employed, so that half of the participants in both switch and no switch conditions used their right hand during the first half of the session, and the remaining half of the participants used their left hand. To insure that participants used the hand assigned to them for a given session half, they wore socks on the unassigned hand. Participants received a 5-minute break between session halves. Thirty-two participants were tested individually in a single session.

We found that, as in Experiment 1, accuracy declined across session halves overall (see Table 1), even though the right hand was used on half of the trials and half of the participants switched from one hand to another halfway through the session. This finding implies that the effect of fatigue is not limited to peripheral motoric processes involving the specific hands but has some central, cognitive component.

In contrast, as in Experiment 1, we found that total correct response times significantly decreased (i.e., improved) across session halves, even though any effects of within-session repetition priming were eliminated in this experiment. We also found a significant interaction of session half and block: There was a general decline in total response time across the five blocks in the first half of the session but no consistent change across the five blocks in the second half of the session.

The response latency decrease across session halves depended on both switch condition and the hand used in the first half because participants were usually faster with their right hand. Specifically, the decrease in total response time across session halves was greatest for the participants who switched from using their left hand in the first half of the session to using their right hand in the second half of the session, and there was a small increase in total response time across session halves for the participants who switched from right hand to left. Importantly, finding the average latency decrease across session halves comparable overall for the switch condition and for the no-switch condition implies that, like fatigue, any learning responsible for the latency decrease is not limited to peripheral motoric processes involving the specific hands but also has a central, cognitive component.

We also examined component response times, and we found different effects of practice on the different components (see Table 1). Most interesting is the fact that initiation time (the time to enter the first digit, including time to encode the number) showed a significant interaction of session half and block; initiation times generally got faster across the first five blocks but got slower across the second five blocks. This pattern suggests that for the relatively difficult and time-consuming process of encoding, in the second half of the session the effect of fatigue overcame the effect of practice found in the first half. In contrast, there was no consistent change across blocks in either session half for execution time (the average time to enter the second, third, and fourth digits) although there was a decrease across session halves. Finally, there was a strong decrease in conclusion time (the time to press the final "enter" key) not only across session halves but also across blocks in both session halves. This pattern leads to the unintuitive suggestion that for purely motoric processes, there may be no effect of fatigue and only an improvement due to practice.

This experiment, therefore, supports our findings from Experiment 1 of a changing response latency-accuracy tradeoff as a function of practice and indicates that the latency decrease with practice cannot be attributed to within-session repetition priming. Further, this experiment illustrates that fatigue affects the component processes of the data entry task differentially and at different points in time, with fatigue having its largest effect on initiation time, which is the most cognitively demanding component.

In the basic version of the data entry experiment, participants are allowed to use whatever means they wish to remember the number, including subvocal or vocal phonological rehearsal, thus activating the phonological loop of working memory. In a new experiment (Kole, Healy, & Buck-Gengler, in press), we wanted to determine whether articulatory suppression, which would disrupt this means of rehearsal, would thus alter performance on this task. We conducted another variant of the initial Buck-Gengler and Healy (2001) data entry experiment to assess this issue.

Specifically, we repeated the initial experiment with 32 new participants and one important change: Half the participants were in an articulatory suppression group, in which they repeated the word "the" continuously while they typed the digits in both sessions, and the remaining participants were in a silent group in which they entered the digits silently, with no secondary articulatory suppression task.

We found that during training, accuracy decreased for the articulatory suppression group relative to the silent group, but only for numbers presented as words. The same interaction was also significant for total correct response time. In this case, however, response times actually *improved* (i.e., got faster) under articulatory suppression, but now only for numbers presented as numerals. An explanation for this unexpected finding is that under articulatory suppression, participants go directly from the printed numeral to the typing response, without subvocalizing the number. Such a strategy could reduce response time but not affect response accuracy. Thus, considering each presentation format (words or numerals) separately, there was no latency-accuracy tradeoff. However, response latency and accuracy did show complementary patterns: Articulatory suppression had negative effects on accuracy for words and had positive effects on latency for numerals.

There was a change in the response time pattern as a function of training block. Specifically, as in our research on fatigue, all participants showed improvement across training blocks; however responses to words presented under articulatory suppression improved at a faster rate than did the other combinations of presentation format and group (see Table 2). This finding is

presumably due to the fact that across blocks participants under articulatory suppression learned to use a more efficient, non-phonological strategy for the word format.

Table 2

Total Response Time in ms in Experiment by Kole, Healy, and Buck-Gengler (in press) During Training by Group, Presentation Format, and Block

Presentation Format	Block				
	1	2	3	4	5
Articulatory Suppression					
Numeral	2497	2449	2361	2283	2264
Word	3478	3305	3205	3104	2958
Silent					
Numeral	2699	2622	2564	2542	2492
Word	3330	3233	3145	3105	3082

Total response times for old numbers were significantly faster than those for new numbers at test 1 week after training, demonstrating repetition priming. This effect was evident only for numbers presented as words at test, not for those presented as numerals at test. The same pattern was found for initiation time. Further, for initiation times, there was an advantage at test for numbers presented as words during training, as found by Buck-Gengler and Healy (2001), resulting in an interaction of presentation format by format continuity for the old items at test. Thus, whenever participants practiced during training typing a given four-digit number presented as a sequence of words, they were faster at test 1 week later than when they practiced during training typing the same number presented as a sequence of numerals, no matter what was the presentation format of the number at test. Importantly, the advantage for numbers presented as words at training was significant for the silent group, but not for the articulatory suppression group. This finding implies that at least some of the advantage for old numbers presented as words at training might be due to subvocalization or use of the articulatory loop.

Initiation times also yielded a surprising main effect of group. Participants who completed the task in the articulatory suppression group initiated trials significantly *faster* at test than did participants who completed the task in the silent group. In contrast, when we examined execution time, we found faster times for the silent group than for the articulatory suppression group, with the *disadvantage* for articulatory suppression greater for words than for numerals. Hence, once again, we found a different pattern of results for the different response time measures. In this case, the pattern seems to imply that the response for the initial digit of the four-digit number can be based on visual input alone, without any phonological code, but that the responses for the subsequent digits do seem to rely on phonological coding, presumably in order to maintain those digits in working memory. These results underline our findings from and our interpretation of the fatigue research that the various response time components reflect different underlying processes. Further, these results suggest that the effects of adding a secondary task cannot simply be described as lowering performance. The effects are more complex, not always negative, and can even be in some respects performance enhancing.

Ways to Promote Transfer of Training

The second issue to be addressed also involves the effects on performance of adding a secondary task, in this case in an effort to understand ways to promote transfer of training.

We have explored many different laboratory and natural tasks in our investigations of the long-term retention and transfer of knowledge and skills. To summarize many of the results of these studies, we formulated the principle of procedural reinstatement (see, e.g., Clawson, Healy, Ericsson, & Bourne, 2001; Healy et al., 1992, 1993; Jensen & Healy, 1998), which is related to the earlier concepts of transfer appropriate processing (e.g., Kolers & Roediger, 1984; Morris, Bransford, & Franks, 1977) and encoding specificity (Tulving & Thomson, 1973). According to this principle, to optimize long-term retention the procedures (i.e., motoric, perceptual, and cognitive operations) required of participants during training must duplicate the procedures used at the retention test. We found that when tasks met this duplication criterion, performance was highly durable over long delay intervals between training and testing. On the other hand, we also found that for such durable tasks, performance was highly specific to the training procedures and did not generalize well even when only minor changes were made in those procedures (see, e.g., Clawson, King, Healy, & Ericsson, 1995; Healy, Wohldmann, & Bourne, 2002; Rickard, Healy, & Bourne, 1994).

Recently, we addressed this issue with a new task, namely the estimation of short temporal intervals (Parker, Healy, & Bourne, 2000). This task is a component in many everyday situations, such as when a speaker has to estimate how much time is remaining in a talk. Although time estimation has been widely studied in the past, the influences of prior training on this skill have not been fully addressed. Because of participants' prior experience in estimating seconds or minutes, we chose a slightly different fundamental unit of time to obtain a purer assessment of the effects of training the skill of duration estimation, with 1 unit equal to 783 ms.

In our prospective production estimation task, participants practice estimating six specified intervals of time expressed in these arbitrary units. Participants are not told how long a given unit is; instead they learn this information through training with feedback, although they naturally know that a larger number of units corresponds to a longer duration than a smaller number of units. For example, participants might see "After the beep, estimate 32 units." They would wait until they thought that 32 units had passed and then press the space bar. Then, they would receive feedback like the following: "Your estimate was 29 units. The difference is -3 units." During training, the intervals were presented in six blocks of six trials.

In our first experiment in this series, after training, participants were immediately given a transfer test with no feedback on three types of intervals. Some of the transfer intervals were repetitions of the actual intervals used during training; some were outside of the practiced range, and others were new within the practiced range. Based on procedural reinstatement, we expected that participants would be most accurate during the transfer test on the actual practiced intervals and least accurate on intervals outside the practiced range. Surprisingly, however, there was no consistent effect of interval type.

These findings concerning transfer interval can be understood in terms of the strategy participants used to estimate intervals. According to retrospective self-reports, almost all participants used some method of counting in order to estimate the durations. Thus, it appeared that participants were using a counting strategy that was highly generalizable to other intervals and independent of the actual estimated intervals, although naturally dependent on the fixed size of the fundamental time unit.

We assumed that counting would rely on some method of articulation in order to maintain an accurate representation of the number of units and the pacing between those units. To test this hypothesis, we conducted a second experiment that included three conditions that varied in the articulation required. The no-secondary-task condition was like that used in our first experiment with no concomitant articulation required. The letter condition required participants simply to repeat continuously a random letter given to them at the beginning of each trial, and the alphabet condition required them to recite aloud the alphabet backward by every third letter also beginning from a random letter given to them at the start of each trial. For example, participants in the alphabet condition might receive the cue "J" and then while estimating the specified interval of 32 units, would say "j, g, d, a, x, u, etc." We expected the counting strategy to be disrupted by this relatively difficult secondary task. No secondary tasks were used during the transfer phase in any of these training conditions. The alphabet task used here requires keeping track of where one is in a sequence of events, and this skill is a component of many everyday situations, such as when a speaker needs to keep track of what he or she has already said and what has to be said next.

We found a large effect of training condition on accuracy during training, with much worse performance for alphabet training relative to training with no secondary task. This finding was consistent with our prediction that this condition would disrupt the usual counting strategy. We also found that performance during transfer was worse in the alphabet condition than in the condition with no secondary task. This finding suggests that the strategies acquired during training were specific to the secondary task, which was not present in any condition during the transfer test. In support of this suggestion, we found that performance on the transfer test was worse than that on the last block of training for both the alphabet and letter conditions (each of which involved a secondary task during training but not during transfer) but not for the condition with no secondary task. Removing the secondary task actually hurt performance for those participants trained with the secondary task. This is a unique finding that is surprising in many respects but consistent with our suggestion concerning the specificity of the strategies learned.

As in Experiment 1, in Experiment 2 we found no support for the prediction that accuracy would be better overall for actual practiced intervals than for new intervals and better for new intervals within the practiced range than for those outside the practiced range. Thus, these two experiments demonstrated that improvement in estimating intervals during training was not specific to the training intervals. On the other hand, we also found that improvement in estimating intervals during training was specific to the secondary task employed. Thus, participants were worse during transfer than at the end of training when the secondary task used during training was not employed during transfer. These findings can be understood within the procedural reinstatement framework on the assumption that the same procedures, based on the fixed fundamental time unit, are used for the different target intervals, but the procedures differ as a function of the secondary task. That is, it appears that participants used a generalizable counting strategy even with the secondary alphabet task. However, the counting strategy was not the same for all conditions but rather depended on the secondary task.

The aim of a new experiment (Healy, Wohldmann, Parker, & Bourne, 2002) was to provide a confirmation and extension of our unexpected findings concerning secondary task effects and to investigate the long-term retention of the duration estimation skill. We examined training of duration estimation skill as well as retraining after a 1-week delay. In both sessions, there were six blocks of trials with feedback, and each block contained six different intervals to estimate.

The intervals we used included three that are relatively low (below 30 s) and three that are relatively high (above 30 s but less than 1 min).

We tested 48 participants in this experiment. There were two different conditions – no switch and switch, depending on whether participants did or did not perform the same task during training and retraining. During both sessions, half of the participants in each condition performed the alphabet task, and the remaining half performed no secondary task. In the previous research, testing always occurred without a secondary task, but retraining occurred with the alphabet task for half of the participants in the present experiment.

We examined two measures of error in this experiment, the proportional absolute error and the proportional relative error. The proportional absolute error is defined as the absolute (i.e., unsigned) difference between the estimated interval and the specified interval divided by the specified interval. This measure gives us a normalized assessment of error magnitude. The proportional relative error is defined as the signed difference between the estimated interval and the specified interval divided by the specified interval. This measure is just like the other one but uses signed values rather than absolute values. It provides us an index of response bias. When the estimated interval is longer than the specified interval, there is a positive bias, by this index, and when the estimated interval is shorter than the specified interval, there is a negative bias.

For proportional *absolute* error, we found significant improvement across blocks, especially across the first two blocks, reflecting the effects of practice, and the improvement was greater in Week 1 than in Week 2. For proportional *relative* error, we found that improvement in estimating intervals reflected a decreasing bias to make estimates toward the central tendency of the practiced range. Thus, there was a negative bias for the intervals of high magnitude and, to a lesser degree, a positive bias for the intervals of low magnitude, especially at the beginning of practice.

Of great interest are the changes that occur between the end of training and the beginning of retraining 1 week later when half of the participants switch tasks. We conducted an analysis comparing performance on Week 1 Block 6 to that on Week 2 Block 1. In terms of proportional absolute error, we found that participants who did not switch tasks showed no change in error across blocks despite the 1-week delay separating the blocks. This finding is consistent with our procedural reinstatement framework and demonstrates remarkable skill durability. Also consistent with the procedural reinstatement framework is the dramatic increase in error across blocks found for participants who did switch tasks, demonstrating remarkable skill specificity.

The decline across blocks as a result of the switch in tasks occurred for both switching directions, although the increase was greater when participants switched from the no alphabet to the alphabet task. When participants switched from the alphabet to the no alphabet task, the level of performance during Week 2 Block 1 was equivalent to that during Week 1 Block 1 for the participants who switched from the no alphabet task to the alphabet task (see Table 3). This result indicates that despite the great improvement from Block 1 to Block 6 in Week 1 for the group who switched from the alphabet to the no alphabet task, that group did not benefit at all from that improvement when it switched from the more difficult alphabet task in Week 1 to the simpler no alphabet task in Week 2. A similar set of results was found by comparing performance on Week 2 Block 1 for the group of participants who switched from the no alphabet task to the alphabet task with performance on Week 1 Block 1 for the participants who switched from the alphabet task to the no alphabet task.

Nevertheless, by comparing the last block of training during Week 2 with the first block of training during Week 2, it is clear that improvement occurred during retraining for all four

groups, especially the groups who switched to a different task during retraining. These groups ultimately reached the same level of performance on their tasks as the two groups who did not switch tasks (see Table 3).

Table 3

Mean Proportional Absolute Error in Experiment by Healy, Wohldmann, Parker, and Bourne (2002) During Critical Blocks

Condition	Week and Block			
	Wk 1 Bl 1	Wk 1 Bl 6	Wk 2 Bl 1	Wk 2 Bl 6
Switch A-N	.306	.188	.216	.138
Switch N-A	.218	.100	.273	.190
No Switch A-A	.303	.185	.185	.172
No Switch N-N	.214	.147	.148	.123

Note. Wk = Week, Bl = Block, A = Alphabet, N = No Alphabet

A similar analysis for the proportional relative error compared Week 1 Block 6 to Week 2 Block 1 to examine performance surrounding the switch in tasks when it occurred. This analysis revealed that the decline in performance for the switch condition relative to the no switch condition at the start of Week 2 was due to an increase in positive bias for low intervals and an even greater increase in negative bias for high intervals.

In summary, this experiment demonstrated clearly that improvement of the time estimation skill during training was completely specific to the task used during training. Participants who were trained with the alphabet task as well as those who were not given that task showed considerable improvement during training. After a 1-week delay participants from those two groups showed perfect retention for the estimation skill that they had acquired if they were in the no switch condition. However, when participants from the switch condition returned 1 week later, they showed an increase in errors, to a large extent when they trained on the easier no alphabet task and retrained on the harder alphabet task but even to some extent when they trained on the harder alphabet task and retrained on the easier no alphabet task. In fact, for both switch groups performance at the start of retraining was equivalent to that for the same task at the start of training. Thus, we found that when the tasks were the same there was no forgetting across the 1-week delay and when the tasks were changed there was no transfer across the same delay. In our current research we are exploring ideas concerning methods that may overcome the lack of transfer and promote generalizability of training for duration estimation.

In theoretical terms, the present results are consistent with the procedural reinstatement principle (e.g., Healy et al., 1992, 1993), assuming that the procedures used are different in the two tasks because duration estimation is integrated with the differing demands in the two tasks. Participants may use a counting strategy in both tasks, but that strategy might involve number counting in the no alphabet task and letter counting in the alphabet task. That is, in the alphabet task, the requirement to say the alphabet backwards by three's becomes integral with the requirement to estimate durations.

In practical terms, these findings highlight the importance of training individuals on the same operations that they will use subsequently. If the operations are changed, then they will start at

square one and training will have been totally useless. Thus, for example, in military or certain industrial applications, if a simulator is used for training, the training may not be worthwhile at all, even if the simulator seems generally realistic, if the full set of operations required in the simulator are different from those required in the field. That is, even if performance in the simulator and performance in the field require the same primary operations, transfer might be limited or even absent if the simulator demands different secondary operations from those demanded in the field. For example, if the simulator lacks specific operations required in the field – such as keeping track of where one is in certain sequences – then training in the simulator may not transfer to the field situation. Therefore, it is important to take into consideration the complete set of field task requirements when developing a training simulator that will effectively prepare individuals for performance in the field.

Conclusions

To conclude, we review the two issues addressed, our findings concerning these issues, and their implications for training. The first issue concerns initiating and executing response components under fatigue produced by prolonged work. We found an increasing response latency-accuracy tradeoff as a function of practice in a data entry task. As training progressed, participants became less and less accurate but generally faster and faster. Further, we found that fatigue affected the component processes of the data entry task differentially and at different points in time during training, with the relatively difficult encoding component (reflected in initiation times) slowing down earlier during training than the other components and the purely motoric concluding component showing no evidence at all of slowing down. In another data entry experiment, we stressed the participants by requiring simultaneous performance of a secondary task. The secondary task was articulatory suppression which eliminates phonological coding of the numbers. Again, we found both harmful and beneficial effects of the stressor. Articulatory suppression reduced typing accuracy for numbers presented as words but reduced overall typing time for numbers presented as numerals. Further, like fatigue, articulatory suppression had different effects on the various response time components. It led to a decrease in initiation time but an increase in execution time, suggesting that encoding can be based on visual input alone, without any phonological code, but that response execution relies on phonological coding, presumably in order to store the digits after the first in working memory. The implications of these findings for training are that adding stressors to a training regime could be harmful or useful depending on what aspects of the task are most crucial.

The second issue addressed concerns ways to promote transfer of training. We used a duration estimation task to study this issue and again looked at the effects of training with a secondary articulatory suppression task. We found that learning was highly specific to the conditions of training, so that participants learned how to estimate durations in a way that critically depended on whether or not they were required to perform a secondary task. The implication for training is that to be effective, training must incorporate the complete set of transfer task requirements, including any secondary task requirements imposed.

REFERENCES

- Buck-Gengler, C. J., & Healy, A. F. (2001). Processes underlying long-term repetition priming in digit data entry. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 879-888.
- Clawson, D. M., Healy, A. F., Ericsson, K. A., & Bourne, L. E., Jr. (2001). Retention and transfer of Morse code reception skill by novices: Part-whole training. *Journal of Experimental Psychology: Applied*, 7, 129-142.
- Clawson, D. M., King, C. L., Healy, A. F., & Ericsson, K. A. (1995). Training and retention of the classic Stroop task: Specificity of practice effects. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 234-254). Thousand Oaks, CA: Sage.
- Fendrich, D. W., Healy, A. F., & Bourne, L. E., Jr. (1991). Long-term repetition effects for motoric and perceptual procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 137-151.
- Healy, A. F., & Bourne, L. E., Jr. (Eds.). (1995). *Learning and memory of knowledge and skills: Durability and specificity*. Thousand Oaks, CA: Sage.
- Healy, A. F., Buck-Gengler, C. J., Kole, J. A., & Bourne, L. E., Jr. (2001, November). *Effects of prolonged work on data entry speed and accuracy*. Paper presented at the 42nd annual meeting of the Psychonomic Society, Orlando, FL.
- Healy, A. F., Clawson, D. M., McNamara, D. S., Marmie, W. R., Schneider, V. I., Rickard, T. C., Crutcher, R. J., King, C., Ericsson, K. A., & Bourne, L. E., Jr. (1993). The long-term retention of knowledge and skills. In D. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 30, pp. 135-164). New York: Academic Press.
- Healy, A. F., Fendrich, D. W., Crutcher, R. J., Wittman, W. T., Gesi, A. T., Ericsson, K. A., & Bourne, L. E., Jr. (1992). The long-term retention of skills. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes, Volume 2* (pp. 87-118). Hillsdale, NJ: Erlbaum.
- Healy, A. F., Kole, J. A., Buck-Gengler, C. J., & Bourne, L. E., Jr. (2004). Effects of prolonged work on data entry speed and accuracy. *Journal of Experimental Psychology: Applied*, 10, 188-199.
- Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2002, November). *Specificity of hand-eye coordination: Training and transfer effects*. Paper presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.

- Healy, A. F., Wohldmann, E. L., Parker, J. T., & Bourne, L. E., Jr. (2002). *Training, retention, and transfer in duration estimation: The procedural reinstatement framework*. Manuscript submitted for publication.
- Jensen, M. B., & Healy, A. F. (1998). Retention of procedural and declarative information from the *Colorado Drivers' Manual*. In M. J. Intons-Peterson & D. Best (Eds.), *Memory distortions and their prevention* (pp. 113-124). Mahwah, NJ: Erlbaum.
- Kole, J. A., Healy, A. F., & Buck-Gengler, C. J. (in press). Does number data entry rely on the phonological loop? *Memory*.
- Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, 23, 425-449.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519-533.
- Parker, J. T., Healy, A. F., & Bourne, L. E., Jr. (2000, November). *Training and transfer of duration estimation: The procedural reinstatement framework*. Paper presented at the 41st annual meeting of the Psychonomic Society, New Orleans, LA.
- Rickard, T. C., Healy, A. F., & Bourne, L. E., Jr. (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1139-1153.
- Schneider, V. I., Healy, A. F., & Bourne, L. E., Jr. (1998). Contextual interference effects in foreign language vocabulary acquisition and retention. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 77-90). Mahwah, NJ: Erlbaum.
- Schneider, V. I., Healy, A. F., & Bourne, L. E., Jr. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, 46, 419-440.
- Schneider, V. I., Healy, A. F., Ericsson, K. A., & Bourne, L. E., Jr. (1995). The effects of contextual interference on the acquisition and retention of logical rules. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 95-131). Thousand Oaks, CA: Sage.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.