

The Effects of Training Under Working Memory Load on Performance During Testing

Mary J. Guhl

Department of Psychology and Neuroscience

Departmental Honors Thesis

University of Colorado at Boulder

April 2, 2014

Thesis Advisor:

Alice F. Healy, Department of Psychology and Neuroscience

Committee Members:

Richard Olson, Department of Psychology and Neuroscience

Jay Ellis, Program for Writing and Rhetoric

Acknowledgements:

James A. Kole for help in designing the experiments, in writing the computer programs for subject testing, and in data tabulation and analysis.

Nathaniel D. Chapman for his support and ideas.

This study was taken in part from a research proposal by Healy, A.F. and Kole, J.A.

Abstract

In 2 experiments, subjects trained in a data entry task, typing one 4-digit number at a time. At training subjects either typed the numbers immediately after they appeared or typed the previous number while holding the current number in working memory. In Experiment 2, either nothing or a physical space was inserted between the second and third digits. Subjects completed a 5-min distractor task after training and then completed the test, which consisted of the same data entry task but with all subjects typing the 4-digit numbers immediately as they appeared. The results indicated little benefit for training using working memory on performance at test; however, a component analysis of each keystroke revealed different strategies for each training condition. The results from Experiment 1 suggested an advantage at test for typing numbers that had been presented during training, but only for subjects who were trained to type the numbers as they appeared, and only for the third keystroke. Experiment 2 showed no advantage overall for having a space between the numbers, and analyses for each keystroke revealed that the extra working memory requirement during training slowed entering the last 3 digits of the number and the concluding keystroke at test.

The Effects of Training Under Working Memory Load on Performance During Testing

Working memory is a short-term memory with limited capacity that is flexible based on the task at hand. For example, consider a person remembering a grocery list that he or she forgot to write down. As the person finds a food item in the grocery store, he or she discards it from working memory, never to be remembered again. Suppose the person remembers an item that he or she initially forgot to include in the mental list. The person adds the new item to the working memory store until he or she finds the new item and then discards it from the mental list. The items in working memory change according to which task a person selectively attends. The purpose of the current study is to test whether having to use working memory during a training session affects performance on a test.

Principles of Training

Researchers have studied training in order to see changes in performance over time, and to determine if training transfers across different circumstances. Several studies have manipulated the type of training, the stimuli presented during training, the length of time between training and testing, and the test itself. However, an explanation of the neural processes thought to underlie training is in order.

Recent studies in neuroscience have found possible neural processes to explain why training is effective. Neuroscientists think that when a well-established memory is retrieved, it is vulnerable to degradation by molecular processes. Forming a memory increases the strength of a synapse (the connection between two neurons) and all of the important activities that help store the memory. By default, neurons want to go back to their base state, with the help of protein degradation processes, which would cancel out the increased activity that formed the memory. When a memory is retrieved, the strengthened synapse is vulnerable to the protein degradation

process. However, in a process called reconsolidation, the act of activating the memory also signals the activities that formed the memory to reactivate, hence the synapse stays strong. Some scientists argue that retrieving a memory does not only make the memory vulnerable to degradation, but that it makes the memory vulnerable to modification (Lee, 2009). In other words, it is possible for the memory to disappear if the reconsolidation process is disrupted, but it is also possible to strengthen the memory to make it stronger than it was originally (Lee, 2009).

Reconsolidation relates to training in that each trial of training strengthens the relevant synapses so that they become stronger after each trial (Lee, 2008). Lee found that training animals in a fear context, followed by more training, strengthened the fear memory, but only after the protein degradation processes had started to destabilize the synapse. When the protein degradation processes were blocked artificially, no reconsolidation was necessary and the strength of the synapse remained the same (Lee, 2008). Another study found that training working memory strengthened the connection between different lobes in the brain and that it increased the efficiency with which attentional resources were distributed in short-term memory (Kundu, Sutterer, Emrich, & Postle, 2013). In other words, training to improve working memory strengthened communication between major lobes of the brain, making short-term memory more efficient and creating less demand on the memory system.

In order to determine what type of training is effective, researchers have manipulated the stimuli within the training sessions and tested the amount of transfer to a testing condition. One study found that when subjects trained in a specific condition and switched to a different training condition, there was no transfer of skill from one condition to another (Healy, Wohldmann, Sutton, & Bourne, 2006). The conditions involved training with one or more computer mouse directions reversed, so that subjects had to acclimate to one or more mouse directions going in

the opposite direction than they were familiar. The only transfer found was when the training and testing conditions a subject could encounter had some similarity, so that the processes involved in training were a subset of those involved in testing (Healy et al., 2006).

Another study used a data entry task similar to the task used in the current study, except varying the presentation of the numbers. Buck-Gengler and Healy (2001) used a data entry task to train four-digit numbers in different formats. One format was presented as numerals (1234) and the other format was presented as words (one two three four). In both conditions, subjects typed the numbers using the 10-keypad. The test consisted of a mixture of old and new four-digit numbers, or numbers that had been presented during training and numbers that had not been presented. The format of the numbers also varied, as in the training session. The results were that subjects who trained with numbers presented as numerals typed slower and made more mistakes on the test than subjects who trained with numbers presented as words. Buck-Gengler and Healy (2001) believe that having to process the words during training led to better retention than seeing a number as numerals and typing it using the same format. Varying the presentation of the training stimuli is one way to test training and transfer, but there have been hypotheses put forward for the results seen from training and transfer.

The specificity of training hypothesis states that training on a specific task in a specific way does not create transfer to a different task. The only transfer that occurs is if the test or other task is the same as what was trained. Chein and Morrison (2010) found that training working memory did not transfer to tasks of reasoning or fluid intelligence, which is the ability to solve problems in new situations separate from an individual's current knowledge and experience. Their experiment, in part, supports the specificity of training concept in that improving working memory during training did not improve performance on different tasks based on reasoning.

Healy et al. (2006) found specificity of training effects when a particular reversal condition during training did not transfer to a different reversal condition, but within the same condition subjects became faster. Not only did subjects not display transfer from one condition to another, they showed improvement when the conditions were the same.

In an experiment that also used reversing a direction on a specific plane, subjects were faster when they tested on the same targets as the ones they trained for initiation time (Wohldmann, Healy, & Bourne, 2008). The experimenters reasoned that initiation, the first keystroke in the data entry task, includes formulating a motor plan as well as making the first move, so when subjects were faster on initiation when they trained and tested on the same targets, it suggested that initiation is specific to the targets that were trained (Wohldmann et al., 2008). This explanation is aligned with the specificity of training hypothesis, at least for making the first movement. In the field of neuroscience, Kundu et al. (2013) speculated that transfer from training to testing “may be contingent on the degree to which the training task and transfer task share common neural architectures” (p. 8714). The idea that transfer from training to testing is based on shared parts of the brain suggests a neural basis for the specificity of training hypothesis.

In contrast to specificity of training, another concept that has been described as an explanation of what goes on between training and testing is contextual interference. Contextual interference is the idea that randomizing the order of presented stimuli during training leads to better retention because the processing demand is greater than training the same stimuli over and over in a block. In addition, constantly changing training induces thinking of different strategies for completing the trials, and thus increases the amount of transfer from one task to another. According to contextual interference, subjects who train in a more difficult setting should

perform worse during training, but better during testing because of the beneficial retention and transfer effects. Shea and Morgan (1979) found evidence for the benefit of training in a random condition compared to training in a blocked condition. The subjects who had trained in the random condition were faster for retention and transfer. The experimenters attributed the faster response time from the random condition to being able to initiate and control the information more efficiently (Shea & Morgan, 1979). Lee and Magill (1983) found that what facilitated performance during training was actually harmful to performance during testing. Specifically, subjects in the blocked condition showed more improvement across training trials than subjects in the random condition, but performed worse during the retention task. Lee and Magill (1983) theorized that the memory trace in the blocked condition was not as established as in the random condition, so retention was worse and the blocked type of training hindered performance.

A more recent experiment tested whether adding distractors to a RADAR detection task would affect performance on a test, but the results were mixed (Young, Healy, Gonzalez, Dutt, & Bourne, 2011). The first experiment found a performance decrement at test when subjects trained with distractors relative to when they trained in a less difficult condition. The other two experiments found performance improvements at test when subjects trained with distractors. Young et al. (2011) theorized that the type of distractors used during training could explain the difference between the experiments. If the distractors used the same mode as the training task (both visual), the cognitive capacity was shared and not enough was left to master the target skill. If the distractors used different sensory modes as the training task (visual and auditory), the cognitive capacity was not shared and performance at test was better.

A review article by Wulf and Shea (2002) acknowledged that there may be a benefit to training with random practice, but argued that learning complex skills may benefit more from

reducing the demands of a task. When first learning a complex task, people who are new to the task may actually be hindered by the extra, randomized task stimuli. However, once people master the task, further improvements may be seen when they are introduced to random training (Wulf & Shea, 2002). As with Young et al. (2011), Wulf and Shea (2002) argue that cognitive capacity is limited, especially when learning a new complex task. In sum, there has not been definitive evidence for supporting a certain way to train, but evidence for separate hypotheses hinge on the type of task and test used and whether distractors for training share cognitive capacity with the primary task.

Performance Throughout Training

An indicator that training has had an effect on learning involves seeing a difference throughout the training session in response time and accuracy. Lohse and Healy (2012) found that when subjects trained in a task that could become proceduralized or automatic, response time was maintained from training to test better than when subjects trained in a task that could not become automatic. However, subjects in the non-automatic condition showed more improvement overall across training in response time and accuracy. Another experiment looked at articulatory suppression and vocalization and its effect on the data entry task (Kole, Healy, & Buck-Gengler, 2005). When subjects trained with articulatory suppression (repeating the word “the” over and over while typing), accuracy was worse across training than subjects who trained in silence, but only if the numbers were presented as words. In contrast, subjects who trained with articulatory suppression typed faster than subjects who trained in silence, but only if the numbers were presented as numerals. The reasoning was that when subjects have to suppress their subvocalization after seeing a number, response time becomes faster because the auditory process is not being utilized in encoding the numbers. In the second experiment, the subjects

who trained in silence were faster than those who vocalized the number they saw, but overall response time decreased across training.

In one type of speed-accuracy tradeoff, as a task becomes mastered, accuracy decreases while response speed increases because of fatigue or boredom and as a result of transfer of controlled movements to automatic movements. One study found that response time in the data entry task improved and accuracy deteriorated across training trials and across training sessions (Healy, Kole, Buck-Gengler, & Bourne, 2004). A similar study done by Kole, Healy, and Bourne (2008) found that improvement in response time was only apparent in the first half of the training session but not in the second half of the session. They believed that cognitive factors, such as mental fatigue, boredom, and the inclination to rest, play more of a part in the speed-accuracy tradeoff than motor factors.

Kole et al. (2008) found that introducing cognitive complications (an extra task such as differing the concluding keystroke or determining which half of the four-digit number was bigger and smaller) while subjects were engaged in the data entry task encouraged greater accuracy for the conditions that had cognitive complications. In the third part of the session, accuracy still dropped for all conditions, but the drop was not as severe for the conditions with cognitive complications. The experimenters proposed that introducing cognitive complications to a task that provokes fatigue or boredom can serve as a cognitive antidote to the observed decline in accuracy throughout a training session. In sum, factors within training contribute to performance across training trials: Some depress performance and others enhance performance.

Retention from Training to Testing

A common way to assess the amount of retention from training to testing includes presenting some stimuli more than once during training, and then the test consists of stimuli from

training as well as stimuli not seen before. Ideally, if subjects were successfully trained, they would show better performance on the stimuli that were presented during training than the new stimuli, a concept called repetition priming. Moving a computer mouse to specified targets in the shape of a clock face on the screen proved to be faster for old targets than for new targets, but the difference became smaller after repeated exposure to the new targets during testing (Wohldmann et al., 2008). Buck-Gengler and Healy (2001) found that in the data entry task, old numbers were typed faster than new numbers overall, but accuracy did not differ between old and new numbers. In another data entry task, Kole et al. (2005) found that old numbers were typed faster than new numbers, but only if the numbers were presented as words.

Fendrich, Healy, and Bourne (1991) presented lists of three-digit numbers over three days. On the first two days, five lists were presented once each and five lists were presented five times each, and subjects typed each three-digit number in the lists. On the third day, each old list was presented once, along with 20 new lists, and subjects had to press a key from 1 to 6 to indicate whether they recognized that list or not. The results showed a repetition priming effect for the lists that were presented five times during training in that the recognition response time was faster. There was not a substantial difference between new lists and lists that were presented once. In other words, subjects recognized the lists that were presented five times and showed it by strongly indicating they were aware that they recognized it and typed the response faster. In sum, repetition priming is one way to assess retention from training to testing. If the response to an old stimulus is better than the response to a new stimulus, then training was effective.

Working Memory

Researchers have done many experiments to test working memory and relevant applications such as working memory capacity and transfer from using working memory to a

different task. Miller (1956) first speculated on the capacity of working memory to be about seven items. He proposed that working memory span is solely limited by the number of items or chunks, independent of the amount of information contained in each item or chunk. For example, remembering seven separate single-digit numbers is the same for Miller as remembering seven two-digit numbers. People can still remember seven chunks regardless of the information contained in the chunks. However, Ericsson, Chase, and Faloon (1980) argued that a more reliable capacity for working memory was four items because the person they were training to expand working memory capacity began to chunk items together in threes and fours after single items became too numerous to reliably remember. Cowan (2000) proposed that the observation of a concrete number of items in working memory can be attributed to experimenters only using items that are not already associated with each other. When memorizing unrelated items, subjects have to focus on all of the unrelated items at once in order to chunk them together and retrieve the chunk later (Cowan, 2000). The limited working memory span has not been observed in tasks where the subjects have a lot of experience with the stimuli, meaning the familiar chunks are stored in long-term memory before being tested (Cowan, 2000).

Based on the observation that working memory shares neural connections with other cognitive skills (Kundu et al., 2013), scientists have tested whether increasing working memory capacity transfers to other processes like attention, intelligence, and cognitive ability in general. After Ericsson et al. (1980) trained their subject to remember a large quantity of numbers, they switched to letters during one session. They found that the large capacity the subject acquired for numbers did not transfer to letters, and the subject's capacity dropped back down to approximately six items for letters. Some theories divide working memory into several components or facets: storage and processing, relational integration, and supervision (von

Bastian, Langer, Jäncke, & Oberauer, 2013). Training in working memory tasks did not promote transfer to reasoning tasks, and focusing on training multiple working memory facets was less efficient than training a single working memory facet (von Bastian et al., 2013). Chein and Morrison (2010) found that training working memory led to improvements in cognitive control and reading comprehension, but not reasoning or fluid intelligence. Their reasoning for seeing different results for different cognitive abilities was that the tasks in which they saw significant transfer may have put more demand on attention, and working memory has been highly linked with attention.

An experiment looking at the association between working memory capacity and a test of attention found that working memory capacity was strongly associated with the ability to control attention (Redick & Engle, 2006). Machizawa and Driver (2011) split visual working memory and attention into three components each and found that a certain component in visual working memory was strongly linked to a particular component in attention, and the same for the other components. For example, the filtering component in visual working memory was linked to the conflict component in attention. Both components are arguably measures of executive function and being able to handle distractors (Machizawa & Driver, 2011).

An experiment testing working memory and fluid intelligence found that training working memory transferred to an increase in performance on tests of intelligence (Jaeggi, Buschkuhl, Jonides, & Perrig, 2008). The experimenters argued that intelligence and working memory share the ability to control attention, so that increasing one will increase the other. However, Chooi and Thompson (2012) did not find that training working memory enhanced intelligence. The literature on working memory is not definitive on the effects of working memory training and whether it transfers to other tasks and cognitive abilities.

The Current Study

The current study attempts to provide more evidence that training using working memory transfers to another task. The study does not look at whether working memory training transfers to other complex cognitive abilities, but simply seeks to find any kind of improvement in performance on a similar task that was trained. The idea for the study came from an experiment that tested whether having to remember to respond in a special way to a specific stimulus led to better performance (Bourne, Healy, Bonk, & Buck-Gengler, 2011). Subjects studied a list of names associated with certain colors and then trained. The training consisted of a name that appeared in the center of the screen surrounded by two rings with blocks of colors. The inner ring had four color blocks and was designated as the default location. The outer ring had four color blocks and was the special response location. Subjects associated the color blocks in the default location with the name in the center unless a special response was indicated (the name that appeared before was in all capital letters), then subjects would have to remember to respond in a different way and choose a color block on the outer ring. The results showed that subjects were less accurate in choosing the correct ring for special responses, but more accurate when choosing a color block for special responses. If having to hold responding in a special way in working memory improved performance in Bourne et al. (2011), then would it be the case that using working memory to train improves performance in general?

One hypothesis of the current experiment was that training using working memory would improve performance on a similar task over not using memory at all. Most other working memory experiments have only looked at transfer from working memory to other cognitive abilities like attention and intelligence. The first hypothesis would also put the contextual interference and specificity of training concepts to test. The second hypothesis was that

performance during training would improve, especially in the working memory condition. A concept that the second hypothesis would test would be the speed-accuracy tradeoff. Perhaps there would be no tradeoff if people had to use working memory. Third, training in the working memory condition would improve memory for old relative to new items, as seen in the concept of repetition priming.

Experiment 1

The purpose of Experiment 1 was to use a data entry task to test whether training using working memory in a delayed typing condition would produce benefits on a test. The test used the same typing task without delaying the typing. If working memory training results in better performance, then having to hold a number in working memory, and delaying the typing, should yield higher accuracy and faster response latency at test than not delaying the typing (immediate typing). Experiment 1 also tested the speed-accuracy tradeoff and repetition priming, in the context of using working memory compared to not using working memory, by analyzing performance across training and including new numbers in the test. If any benefit exists for using working memory to train, testing should ideally reveal that using working memory resulted in improvement throughout training and better performance in typing old relative to new numbers at test.

Method

Subjects

The subjects were 48 undergraduate college students from the University of Colorado at Boulder. All subjects were sampled from the introductory psychology pool and compensated with credit for their class. The sample consisted of 26 men and 22 women aged 18 to 22.

Design

The training phase used a 2x5 mixed factorial design and included two independent variables. Independent variables are variables that the experimenter manipulates, such as whether subjects type a number immediately when they see it or delay typing to the next screen in the context of the present experiment. Independent variables are often divided into between-subjects and within-subjects. Between-subjects variables keep all conditions separate: a subject only participates in one condition. Within-subjects variables allow subjects to be involved in more than one condition. For example, one variable in the present experiment was the between-subjects variable of training condition: delayed or immediate. In the immediate condition, subjects simply typed the four-digit number that appeared on the screen and pressed the “enter” key. In the delayed condition, subjects had to remember the four-digit number that appeared on the screen, press “enter,” and then type that number on the next screen while viewing the next four-digit number, thus encouraging the use of short-term memory while training. The second independent variable was the within-subjects variable of training block in which all subjects participated regardless of training condition. The training phase consisted of five blocks with 65 four-digit numbers in each block.

The testing phase was a 2x2 design with two independent variables. The first was the between-subjects variable of training condition: whether subjects trained in the immediate or delayed condition. The second independent variable was the within-subjects variable of item type, or new versus old numbers. The testing phase contained 128 four-digit numbers. Sixty-four were numbers that had appeared in the training phase (old), whereas 64 of the numbers were ones that did not appear in the training phase (new). Each four-digit number during the test that

had appeared during training (old numbers) had been shown five times during training, once in each block.

The dependent variables, or the variables that measure the response observed from manipulating the independent variables, were accuracy and response latency. Accuracy was defined as all or nothing. If subjects typed one number wrong the program counted the whole four-digit number as incorrect or inaccurate. Response latency is the time it takes to type a number, or the time it takes to press a key on the keyboard. By this definition, higher response latency means more time taken to type a number, or a slower response, whereas lower response latency means less time taken, or a faster response. Overall response latency was measured, as well as response latency for each keystroke including the time taken to type the first number (initiation time), the average time taken to type the second, third, and fourth numbers (execution time), and the time taken to press the “enter” key (conclusion time).

Materials

The computers used were iMac OS X version 10.6.8, 2 GHz Intel Core 2 Duo, 1 GB DDR2 SDRAM desktop computers with Apple Pro keyboards. The program used was written in REALbasic. The program used two sets of numbers to ensure that the particular numbers chosen did not have any effect on the results, so that half of the subjects had a different number set than the other half: sets A and B. Each number set used a total of 65 distinct four-digit numbers during training. The program presented each number five times, once in each block. The test had 64 of the 65 numbers from training as well as 64 of the 65 numbers from the other set, so that in the test set A had 64 numbers taken from training and 64 numbers from the training phase in set B, and set B had 64 numbers taken from training and 64 numbers from the training phase in set A. Subjects sat 50 to 75 cm away from the screen, and the font used in the program was 20-point

Courier. Each four-digit number appeared at the same location on the screen above and to the left of the center. Instructions were displayed on the screen when subjects entered the room with the computer in it. The program contained a short practice block consisting of five four-digit numbers, then the five training blocks and the testing block.

In the immediate condition, the last four-digit number in each block was in red font, signifying that the subject should only press the “enter” key and go on to the next block. In the delayed condition, the first number in each block was in red font, signifying that the subject should remember that number and press the “enter” key to go on to the next screen. The red number gave the delayed condition a starting point for memorizing the numbers while not giving the immediate condition an extra four-digit number to type. In sum, both training conditions had five blocks with 65 four-digit numbers shown in each block, but subjects only typed 64 numbers in each block. For both conditions (delayed and immediate) during both training and testing, the four-digit number stayed on the screen until the subject pressed the “enter” key. Before each phase (practice, training, and testing) a tone sounded three times to make the subjects aware that the phase was beginning. After the practice block, a message appeared on the screen (“You have completed the practice block. Please press the space bar.”). After each training block a message appeared on the screen (“You have reached the end of Block N.”), allowing a few seconds of inactivity between each block. At the end of the training phase, a message appeared on the screen (“Please contact the experimenter.”), and at the end of the testing phase a message appeared (“You have completed the experiment.”).

Between the training phase and the testing phase, subjects completed a letter detection distractor task with pen and paper. The task consisted of two passages adapted from a book on infant care. Half of the passages had a solid color font (all black or all red), and half of the

passages had red and black letters alternating throughout the passage. After each passage appeared a comprehension question asking about the content of the passage. Each subject received a packet with the consent form, instructions, demographic questions, a passage followed by a question, and another passage followed by another question. The letter detection task put some time between the training and testing phases and distracted the subjects from thinking about the training they had just finished using a task with different modes: writing and reading letters instead of typing and reading numbers.

Procedure

The experimenter tested the subjects in a lab room with several cubicles, each containing a computer. They were assigned to a condition using a fixed rotation based on when they arrived for the experiment. When the subjects arrived at the lab room, the experimenter asked them to read the consent form and sign it if they agreed to participate. Then the experimenter took each subject into a cubicle with a computer and asked them to read the instructions on the screen.

The instructions described the experiment procedure, including the three components (practice, training, testing), not to type anything when they see the red font except the “enter” key, to use their preferred hand and the 10-digit keypad on the right side of the keyboard, to not go back and fix mistakes but keep going, and to type the numbers as quickly and accurately as possible. Subjects could not see the numbers as they typed, and the delete key was not active. The instructions in the immediate condition told subjects to type the four-digit number that appeared on the screen; for example, if 1234 appeared, the subject would type 1234 and press the “enter” key. The instructions in the delayed condition told subjects to remember the number on the current screen and then type the number on the next screen and remember the number on that screen; for example, if 1234 appeared as the first number in a block, the subject would remember

that number, press the “enter” key, then type that number on the next screen, and remember the number that appeared on the same screen. After the subjects read the instructions, the experimenter entered the cubicle and asked the subjects to repeat back what they had to do. Once the experimenter was satisfied with their response and answered any questions, the experimenter entered a key combination to start the practice block.

When the subjects finished with the practice and training blocks, they received the letter detection packet and were told that they would be doing a different task for a few minutes. They read and signed the consent form and answered the demographic questionnaire, which consisted of three questions (“Have you ever participated in a letter detection experiment?” “Was English the first language you spoke?” “What is your gender?”) Then they read the instructions, which told half of them to circle the letter “h” in each passage and the other half to circle the word “the” in each passage. The instructions also told the subjects to not go back if they missed an “h” or “the” and to not go back to the passages for an answer when they got to the comprehension question after the passage. They read the passages and answered the comprehension questions, all of which took 3 to 5 min.

After the distractor task, the experimenter told the subjects that they would be completing the testing phase. For the immediate condition, the experimenter told subjects that they would be doing what they had done before, typing what they see as it appeared on the screen and pressing the “enter” key. The experimenter told subjects in the delayed condition that instead of typing the current number on the next screen, they should type the number they see on the screen as it appeared on that screen. In other words, every subject in both conditions completed the same testing phase, doing exactly what the immediate condition did all along, typing the numbers as they appeared on the screen and pressing the “enter” key.

Results

Mixed factorial ANOVAs were conducted for all measures, which included training condition (immediate versus delayed), test item type (new item versus old item), and block number for training. An analysis was done on overall accuracy for training and testing. Additionally, separate analyses were conducted for keystroke latency, which was separated into initiation (first keystroke), execution (second, third, and fourth keystrokes), and conclusion (fifth keystroke or the “enter” key). Additionally, Fisher’s PLSD post hoc tests were done to find significant differences between each keystroke for the execution time (second, third, and fourth keystrokes). Total response latency is reported only for training because the significant result found for testing was apparent in the component keystroke latencies and is reported separately. All response latencies reported are in seconds, and all significant results have a $p < .05$ threshold.

Training

Accuracy. The only significant finding for accuracy was a main effect for block number, $F(4, 184) = 3.318$, $MSE = 0.002$, $p = .0119$. Interestingly, subjects became less accurate with each block (Block 1, $M = .921$, $SEM = .008$; Block 2, $M = .912$, $SEM = .010$; Block 3, $M = .904$, $SEM = .010$; Block 4, $M = .891$, $SEM = .012$; Block 5, $M = .894$, $SEM = .009$).

Overall response latency. A main effect was found for block number, $F(4, 184) = 21.320$, $MSE = 0.029$, $p < .0001$. Subjects became faster at typing the numbers across blocks overall (Block 1, $M = 2.502$, $SEM = .086$; Block 2, $M = 2.377$, $SEM = .085$; Block 3, $M = 2.339$, $SEM = .080$; Block 4, $M = 2.244$, $SEM = .071$; Block 5, $M = 2.218$, $SEM = .072$). However, Figure 1 shows a significant interaction for block number and training condition, $F(4, 184) = 2.589$, $MSE = 0.029$, $p = .0383$. Subjects in the delayed condition showed more improvement in response latency across blocks than subjects in the immediate condition. The result that subjects

became faster, coupled with the decline in accuracy across blocks, suggests a speed-accuracy tradeoff. As subjects train for longer periods, response latency improves as they learn the task and it becomes more automatic, but accuracy decreases because the task has become automatic and subjects do not pay as much attention to what they are doing.

Initiation. A main effect was found for training condition, $F(1, 46) = 240.804$, $MSE = 0.138$, $p < .0001$. Surprisingly, subjects in the delayed condition typed the first keystroke much faster ($M = .337$, $SEM = .013$) than subjects in the immediate condition ($M = 1.080$, $SEM = .021$), providing a hint that each training condition was using a different strategy to master the task. Another main effect was found for block number, $F(4, 184) = 8.809$, $MSE = 0.012$, $p < .0001$. Overall, subjects typed the first keystroke faster with each block (Block 1, $M = .790$, $SEM = .067$; Block 2, $M = .703$, $SEM = .059$; Block 3, $M = .688$, $SEM = .060$; Block 4, $M = .690$, $SEM = .061$; Block 5, $M = .675$, $SEM = .056$).

Execution. A main effect was found for keystroke, $F(2, 92) = 43.235$, $MSE = 0.015$, $p < .0001$. The third keystroke was the slowest ($M = .319$, $SEM = .007$) for execution time, followed by the second ($M = .234$, $SEM = .004$) and fourth keystroke ($M = .223$, $SEM = .003$). A post hoc test revealed a significant difference between the second and third keystrokes ($p < .0001$) and between the third and fourth keystrokes ($p < .0001$), emphasizing the substantial difference in reaction latency between the third keystroke and the other two and a possible chunking strategy. Figure 2 shows a significant interaction for keystroke and training condition, $F(2, 92) = 11.415$, $MSE = 0.015$, $p < .0001$. Subjects in the immediate condition typed the third keystroke slower ($M = .362$, $SEM = .013$) than subjects in the delayed condition ($M = .275$, $SEM = .004$); otherwise both training conditions had comparable speeds for the second and fourth keystrokes. Lastly, a main effect was found for block number, $F(4, 184) = 7.045$, $MSE = 0.001$, p

$< .0001$. As expected, subjects became faster at typing the second, third, and fourth keystrokes with each block (Block 1, $M = .269$, $SEM = .007$; Block 2, $M = .264$, $SEM = .008$; Block 3, $M = .258$, $SEM = .007$; Block 4, $M = .253$, $SEM = .008$; Block 5, $M = .249$, $SEM = .007$).

Conclusion. An analysis found a main effect for training condition, $F(1, 46) = 48.133$, $MSE = 0.980$, $p < .0001$. Subjects in the delayed condition were a lot slower ($M = 1.134$, $SEM = .059$) on pressing the “enter” key than subjects in the immediate condition ($M = .247$, $SEM = .004$), indicating more evidence for a difference in strategies between the training conditions. Another main effect was found for block number, $F(4, 184) = 4.914$, $MSE = 0.020$, $p = .0009$. Overall, subjects became faster at typing the “enter” key with each block (Block 1, $M = .735$, $SEM = .099$; Block 2, $M = .717$, $SEM = .100$; Block 3, $M = .715$, $SEM = .099$; Block 4, $M = .653$, $SEM = .084$; Block 5, $M = .631$, $SEM = .079$). The top panel of Figure 3 shows a significant interaction for block number and training condition, $F(4, 184) = 3.924$, $MSE = 0.020$, $p = .0044$. Subjects in the immediate condition remained static across blocks, whereas subjects in the delayed condition became faster at typing the “enter” key.

Testing

Accuracy. Nothing significant was found for accuracy during testing, suggesting a possible ceiling effect for the test. Not finding a result for accuracy also suggests that there is no benefit to using working memory during training to increase accuracy on a test.

Initiation. There were no significant effects for initiation time during testing, including training condition and item type, which indicates that having to train using working memory had no effect on performance on the first keystroke for testing. Also, whether a number was new or old had no effect on the first keystroke during testing.

Execution. A main effect was found for keystroke, $F(2, 92) = 36.052$, $MSE = 0.013$, $p < .0001$. The third keystroke was the slowest ($M = .330$, $SEM = .015$) during testing, followed by the second ($M = .219$, $SEM = .005$) and fourth ($M = .200$, $SEM = .004$) keystrokes. A post hoc test showed significant differences between the second and third keystrokes ($p < .0001$) and the third and fourth keystrokes ($p < .0001$), as in the execution time for training. The slower third keystroke during training for the immediate condition and for both training conditions overall during testing provides evidence for the subjects adopting a chunking strategy. There was also a main effect for item type, $F(1, 46) = 17.651$, $MSE = 0.001$, $p = .0001$. New numbers were typed slower ($M = .255$, $SEM = .010$) than old numbers ($M = .244$, $SEM = .008$).

A significant interaction was found for item type and training condition, $F(1, 46) = 5.898$, $MSE = 0.001$, $p = .0191$. Subjects who trained in the immediate condition typed new numbers slower ($M = .268$, $SEM = .017$) than subjects who trained in the delayed condition ($M = .243$, $SEM = .011$), but the immediate condition showed more improvement typing old numbers ($M = .250$, $SEM = .014$) than subjects in the delayed condition ($M = .238$, $SEM = .010$) if seen from new to old. In other words, there is more of a difference between typing new and old numbers in the immediate condition than there is in the delayed condition. However, taking old numbers by themselves, the delayed condition typed old numbers faster than the immediate condition. Figure 4 shows an interaction for keystroke and item type, $F(2, 92) = 7.946$, $MSE = 0.001$, $p = .0007$. Although the second and fourth keystrokes were comparable in speed between new and old numbers, the third keystroke was typed slower for new numbers ($M = .343$, $SEM = .023$) than for old numbers ($M = .317$, $SEM = .019$). Finally, Figure 4 also shows a significant three-way interaction for keystroke, item type, and training condition, $F(2, 92) = 4.576$, $MSE = 0.001$, $p = .0127$. Subjects who trained in the immediate condition typed new numbers slower ($M = .383$,

$SEM = .037$) than old numbers ($M = .338$, $SEM = .031$) for the third keystroke, whereas subjects who trained in the delayed condition did not change substantially between new ($M = .304$, $SEM = .027$) and old ($M = .295$, $SEM = .024$) for the third keystroke.

Conclusion. There was a main effect for training condition, $F(1, 46) = 15.498$, $MSE = 0.009$, $p = .0003$. Subjects that had trained in the delayed condition were slower ($M = .297$, $SEM = .013$) to type the “enter” key than subjects that had trained in the immediate condition ($M = .222$, $SEM = .005$), suggesting that subjects in their respective training conditions kept the same strategy during testing that they adopted during training. Another main effect was found for item type, $F(1, 46) = 4.443$, $MSE = 0.0002446$, $p = .0405$. New items were typed slower ($M = .263$, $SEM = .012$) than old items ($M = .256$, $SEM = .010$).

Discussion

In general, Experiment 1 sought to test whether using working memory during training would lead to better performance on a test. Additionally, retention from training to testing was measured using old and new numbers. An analysis of training by itself revealed a speed-accuracy tradeoff in that accuracy declined and response latency decreased across the five blocks. The result could be attributed to the typing portion of the data entry task becoming more automatic throughout training, switching from conscious executive control to unconscious motor control. If subjects did not have a good grasp of the layout of the 10-keypad at the start of training, they would have had to look at each number on the keypad, using conscious executive control to direct their fingers to move. As the task became familiar, subjects would have been able to relinquish conscious control of the movement to automatic control. The decline in accuracy would be attributed to subjects not paying as much attention to the movement of their fingers because it became automatic. The decrease in response latency would be because the movement

of fingers became automatic. The result could also be attributed to subjects simply being bored and disengaged from the task.

An analysis of testing by itself confirmed the effect of repetition priming. New numbers were typed slower for execution overall and conclusion time, suggesting that training for old numbers did improve typing the old numbers during testing. In addition, subjects typed the third keystroke more slowly for new numbers than old numbers, but there was essentially no difference between new and old for Keystrokes 2 and 4 (see Figure 4). Similar to Keystroke 1 measuring encoding and motor planning (Fendrich et al., 1991), Keystroke 3 could be measuring more encoding as the subjects look at the first two numbers then the second two numbers, resulting in a chunked response.

Contextual interference, as defined by Shea and Morgan (1979), is the idea that training in a condition with randomly presented stimuli leads to better transfer to a test than training in a condition with blocked stimuli because of greater processing demands in random training. If contextual interference could be extended to increasing the demand on working memory, the delayed condition could be thought of as using contextual interference. Ideally, subjects hold two four-digit numbers in working memory: the number from the previous screen so they can type it and the number on the current screen so they can type it on the next screen. The two four-digit numbers would interfere with each other in working memory, creating more processing demand.

During training, subjects in the delayed condition showed more improvement for total response latency across blocks than subjects in the immediate condition (see Figure 1). Subjects in the immediate condition likely had less room for improvement; however, total response latency was not significant for the training conditions. If contextual interference truly had an effect, the delayed condition should have been slower and less accurate overall. Contextual

interference indicates that subjects perform worse during a training session but better during testing because of the difficult training (Shea & Morgan, 1979). The current experiment found no evidence for improved accuracy during testing from the delayed condition. Subjects in the immediate condition showed more improvement in response latency for execution time from new numbers to old numbers than subjects in the delayed condition, providing evidence that contextual interference had no effect for execution response latency. In particular, subjects in the immediate condition typed slower on Keystroke 3 for new numbers than old, whereas there was essentially no difference on Keystroke 3 between new and old numbers for subjects in the delayed condition (see Figure 4).

The result did not support contextual interference because new and old numbers for execution time did not make a difference for subjects in the delayed condition. If contextual interference had an effect, subjects in the delayed condition should have typed the old numbers faster than new numbers, and typed old numbers faster than subjects in the immediate condition. The result could also be interpreted as subjects in their respective training conditions adopted different strategies for the task and kept the same strategies during testing. Subjects in the immediate condition chunked the numbers because they could see them on the screen and refer back to them, and subjects in the delayed condition remembered the numbers in one big chunk and, therefore, did not chunk their response on the next screen. For the second explanation, presenting new versus old numbers would not make a difference for the strategy that subjects in the delayed condition adopted because they did not take as much time to encode each number, as seen by the lack of observation of chunking in the delayed condition.

The hypothesis that using working memory during training would improve performance on a test had little evidence. Subjects in the delayed condition did not have greater accuracy or

faster overall response latency on the test than subjects in the immediate condition. The results show that there was no statistically significant difference between the training conditions for initiation and execution for the test. During the test, when subjects in the delayed condition did not have to use their working memory, they possibly started to adopt the same strategy as subjects in the immediate condition: unconsciously taking the time to encode and plan the movement for the number before typing it.

Subjects in the immediate condition typed the third keystroke slower for new numbers than old numbers, whereas there was essentially no difference for new and old numbers in the third keystroke for subjects in the delayed condition (see Figure 4). The difference between immediate and delayed for the third keystroke could suggest better retention for training in the immediate condition, and when all keystrokes are taken into account, the delayed condition typed slightly slower on the test overall. An explanation for why the delayed condition did not perform better on the test than the immediate condition could be the strategy that subjects in the delayed condition adopted during training and transferring it to the test. Looking at all four numbers as one big chunk did not help performance in the test for the delayed condition.

Experiment 1 had a major flaw that could have affected the results. The four-digit number stayed on the screen until the subjects pressed the “enter” key. In the delayed condition, therefore, the only obstacle between seeing the number and typing it on the next screen was pressing the “enter” key. Subjects were not forced to use their working memory extensively, and nothing interfered with the number in working memory and the current number. They probably did not even look at the current number until they had typed the number from the previous screen.

Multiple results were found revolving around the first and third keystrokes, indicating that subjects possibly chunked the four-digit number mentally in order to remember it so that it could be typed easily. Chunking was especially apparent in the immediate condition (see Figure 2), where subjects could have looked at the first two digits and typed them, then looked at the second two digits and typed them. They also could have looked at the four-digit number as a whole, mentally chunked it into two chunks, and typed it as two chunks because they had in mind two chunks. Subjects in the delayed condition did not show prominent chunking, possibly because when a new screen appeared, they immediately started typing the number from the previous screen, then used the conclusion time to encode and develop a motor plan for the number on the present screen before pressing “enter” (see top panel, Figure 3). The delayed strategy would account for the faster initiation time, little difference between the second through fourth keystrokes, and the slower conclusion time. However, a question remains whether subjects in the immediate condition mentally chunked the numbers. If they did mentally chunk the numbers, then the question could be whether mentally chunking the numbers helped them on the test. Extending the question of chunking could include whether forcing the subjects in the delayed condition to think of the numbers as chunks would help them on the test as well.

Experiment 2

The goal of Experiment 2 was to address the issues of Experiment 1 as well as explore the chunking found in Experiment 1. Experiment 2 also tested the effects of training using working memory and subsequent retention to testing, as in Experiment 1. Instead of the four-digit numbers remaining on the screen until subjects pressed the “enter” key, both training conditions had a time limit that the numbers appeared on the screen. Shortening the time that the numbers stay on the screen should make it so subjects in the delayed condition have to actually

use their working memory and update it with new numbers, but without substantially affecting performance in the immediate condition.

Experiment 1 found that subjects typed the first and third keystroke more slowly, indicating chunking. Chunking is a mnemonic device frequently used to enhance working memory capacity. When a list becomes too long, people will automatically start chunking the items in the list into groups in order to remember the list. Miller (1956) proposed that working memory span could be increased by combining items to be remembered into chunks, increasing the size of each chunk until all the items could be remembered. However, Ericsson et al. (1980) argued that using chunks of information did not expand working memory capacity. They proposed that the supposed increase in working memory capacity was due to the subject using chunks to create associations in long-term memory, not due to practice in expanding the capacity of working memory.

Using a task with spatial stimuli, Ridgeway (2006) found that when subjects adopted a chunking strategy, accuracy improved on the task. Similar to Ericsson et al. (1980), Ridgeway argued that although people can easily remember five or six items, the actual average working memory capacity smaller because of the unconscious use of chunking. One study explored the difference between presenting stimuli as chunks and requesting responses in chunks (O'Shea & Clegg, 2006). In the first two experiments, the researchers found that presenting stimuli in chunks increased proportion of correct items relative to not presenting the stimuli as chunks, requesting a response with chunks, and requesting a response without chunks. The third experiment, however, revealed that responding with chunking may improve performance if the stimuli were presented as random chunks (varying the number of items in each chunk), suggesting that responding with chunking may indeed play a role in improving performance in

specific conditions. An explanation for the results in the third experiment is that having to respond with chunking acts as another way to rehearse the items in the list and group them in a way that is easy to remember (O'Shea & Clegg, 2006).

Based on what previous studies have found regarding the advantage to chunking, Experiment 2 sought to expand the chunking result found in Experiment 1 by testing whether forcing subjects to chunk improved performance on the test. Consequently, Experiment 2 introduced a space between the second and third digits for half of the subjects during training, thereby creating two chunks with two numbers in each chunk. Forcing subjects to divide the four-digit numbers into chunks should improve performance throughout training and on the test.

Method

Subjects

Forty-eight subjects were taken from the University of Colorado at Boulder's introduction to psychology pool. They were compensated with credit for their class. The average age was 18 to 22 and the sample consisted of 24 men and 24 women. The experimenter tested an additional eight subjects but their data was excluded because they did not follow instructions or pressed the "enter" key through entire training blocks.

Design

A 2x2x5 mixed factorial design was used for the training phase including three independent variables. Two of the independent variables were the same as Experiment 1: training condition (delayed or immediate) and training block (1 through 5). Another between-subjects independent variable was introduced based on the results from the first experiment: training gap presence (space or no space). Half of the subjects, regardless of training condition, had

approximately one centimeter of space between the second and third numbers in each four-digit number during training.

As in Experiment 1, the testing phase was a 2x2x2 design with the added third variable: training condition, item type or new versus old numbers, and whether subjects trained with a space between the numbers or not. The test was constructed exactly as it was in Experiment 1. The subjects who saw the space between numbers during training did not see it during testing.

The dependent variables were the same as Experiment 1: accuracy and response latency. They were measured overall and across the blocks of training and testing. Response latency for each keystroke was also measured as in Experiment 1.

Materials

The same computers used in Experiment 1 were used in Experiment 2, except the keyboards were switched to Empirisoft Millisecond Accurate Keyboards. The same program, REALbasic, was used with two sets of numbers. Some minor modifications were made. The font was enlarged to 30 point, and instead of the numbers staying on the screen until the subjects pressed the “enter” key, each four-digit number stayed on the screen for 2,000 ms for both training conditions to address the problem of subjects not really using their working memory in the delayed condition in Experiment 1. Instructions did not change except to reflect a space between the second and third numbers (for example, 12 34) in the chunking conditions. The layout of the program was the same as Experiment 1 with a short practice block, five training blocks, and testing, and with the same amount of numbers in each phase as in Experiment 1.

As in Experiment 1, subjects completed a letter detection task between the training and testing phases. The task was the same except the letters varied by capital and lowercase instead of the colors red and black.

Procedure

The procedure was the same as in Experiment 1. Subjects were assigned to a condition using a fixed rotation based on when they arrived. They completed the practice block and training phase, did the letter detection distractor task, and completed the testing phase. No mention of the space between the numbers in the chunking conditions was ever referenced by the experimenter or the instructions in the program. The training conditions completed the same test, typing the number immediately after it appeared and with no space between the second and third numbers.

Results

Mixed factorial ANOVAs were performed for all measures including the factors of training condition, training gap presence, test item type for testing, and block number for training. Accuracy and keystroke latency were also included as measures, with keystroke latency broken down into initiation, execution, and conclusion. Initiation is the first keystroke, execution is the average of the second, third, and fourth keystrokes, and conclusion is the fifth keystroke or the “enter” key. Additional post hoc analyses conducted were Fisher’s PLSD to identify significant paired comparisons within the execution time, or second, third, and fourth keystrokes. Overall response latency was only reported for testing and not training because the only significant result for training was only significant because of one of the keystrokes, which is reported separately. A significant effect was also found for overall response latency for training condition and training gap presence during testing, but is not reported because one of the component keystrokes made the effect significant and is reported within that component’s section. All response latencies reported are in seconds. All significant effects ($p < .05$) are reported.

Training

Accuracy. A main effect was found for training condition, $F(1, 44) = 48.743$, $MSE = 0.068$, $p < .0001$. As expected, subjects in the immediate condition were much more accurate ($M = .910$, $SEM = .005$) than subjects in the delayed condition ($M = .675$, $SEM = .016$). Block number was also significant for accuracy overall, $F(4, 176) = 9.730$, $MSE = 0.004$, $p < .0001$. Overall, subjects showed improvement in accuracy across all blocks (Block 1, $M = .754$, $SEM = .031$; Block 2, $M = .771$, $SEM = .028$; Block 3, $M = .812$, $SEM = .024$; Block 4, $M = .812$, $SEM = .023$; Block 5, $M = .815$, $SEM = .022$). Figure 5 shows a significant interaction for block number and training condition, $F(4, 176) = 14.751$, $MSE = 0.004$, $p < .0001$. Subjects in the immediate condition did not get any more accurate across blocks -- in fact they became slightly less accurate -- whereas subjects in the delayed condition became more accurate. A speed-accuracy tradeoff for subjects in the immediate condition could have occurred, and suggests fatigue or boredom. However, subjects in the delayed condition became more accurate and faster across blocks, suggesting the task was difficult enough to keep those subjects engaged in the task.

Initiation. A main effect was found for training condition, $F(1, 44) = 25.009$, $MSE = 1.091$, $p < .0001$. Subjects in the delayed condition were faster on the first keystroke ($M = .652$, $SEM = .045$) than subjects in the immediate condition ($M = 1.326$, $SEM = .048$), suggesting the adoption of different strategies for each training condition. A main effect was also found for block number, $F(4, 176) = 6.931$, $MSE = 0.062$, $p < .0001$. Overall, subjects typed faster over time in each block (Block 1, $M = 1.123$, $SEM = .085$; Block 2, $M = 1.040$, $SEM = .092$; Block 3, $M = .974$, $SEM = .095$; Block 4, $M = .916$, $SEM = .082$; Block 5, $M = .892$, $SEM = .086$).

Execution. There was a main effect for training gap presence, $F(1, 44) = 5.672$, $MSE = 0.019$, $p = .0216$. Subjects overall were slower to type keystrokes 2 through 4 with a gap ($M =$

.269, $SEM = .006$) than without a gap ($M = .245$, $SEM = .003$). There was a main effect found for keystroke, $F(2, 88) = 81.061$, $MSE = 0.013$, $p < .0001$. For execution, the slowest keystroke was the third ($M = .332$, $SEM = .006$), followed by the second keystroke ($M = .226$, $SEM = .003$), and the fourth keystroke was the fastest ($M = .213$, $SEM = .002$). A post hoc test revealed significant mean differences between Keystrokes 2 and 3 ($p < .0001$) and between Keystrokes 3 and 4 ($p < .0001$), suggesting a chunking strategy in which the third keystroke contains encoding time and the fourth keystroke becomes shortened as subjects slide into the conclusion keystroke. A main effect was found for block number, $F(4, 176) = 15.461$, $MSE = 0.001$, $p < .0001$. Overall, subjects typed keystrokes two through four faster with each block (Block 1, $M = .269$, $SEM = .008$; Block 2, $M = .265$, $SEM = .008$; Block 3, $M = .258$, $SEM = .007$; Block 4, $M = .248$, $SEM = .006$; Block 5, $M = .245$, $SEM = .006$).

Figure 6 shows a significant interaction for keystroke and training gap presence, $F(2, 88) = 13.736$, $MSE = 0.013$, $p < .0001$. Subjects who were presented with the gap were slower on Keystroke 3 ($M = .376$, $SEM = .010$) than subjects who were presented without the gap ($M = .289$, $SEM = .006$), suggesting the space enhanced the chunking strategy. Figure 7 shows a significant interaction for block number and training condition, $F(4, 176) = 5.140$, $MSE = 0.001$, $p = .0006$. Subjects in the delayed condition were slower, especially in the first block, than subjects in the immediate condition. Figure 8 shows a significant interaction for keystroke and block number, $F(8, 352) = 4.952$, $MSE = 0.001$, $p < .0001$. While the second and fourth keystrokes remained relatively static across blocks, the third keystroke was typed significantly faster across blocks but remained slower than the second and fourth keystrokes. Finally, Figure 8 also shows a significant three-way interaction for keystroke, block number, and training condition, $F(8, 352) = 5.758$, $MSE = 0.001$, $p < .0001$. Subjects in the delayed condition typed

the third keystroke slower than subjects in the immediate condition, especially in the first block. However, subjects in the delayed condition showed more improvement for typing the third keystroke across blocks, but subjects in the immediate condition became slower at typing the third keystroke in the second block than in the first block and then gradually improved.

Conclusion. There was a main effect for training condition, $F(1, 44) = 16.336$, $MSE = 0.681$, $p = .0002$. Subjects in the delayed condition were slower on pressing the “enter” key ($M = .673$, $SEM = .053$) than subjects in the immediate condition ($M = .243$, $SEM = .003$). Another main effect was found for block number, $F(4, 176) = 2.659$, $MSE = 0.030$, $p = .0344$. Overall, subjects became slower on pressing the “enter” key over time and across blocks (Block 1, $M = .397$, $SEM = .049$; Block 2, $M = .435$, $SEM = .059$; Block 3, $M = .482$, $SEM = .073$; Block 4, $M = .484$, $SEM = .073$; Block 5, $M = .492$, $SEM = .077$), which is the opposite result that the speed-accuracy tradeoff predicts. However, the bottom panel of Figure 3 shows an interaction for block number and training condition, $F(4, 176) = 3.197$, $MSE = 0.030$, $p = .0145$. Across blocks, subjects in the immediate condition pressed the “enter” key at about the same speed or slightly faster, which does support the speed-accuracy tradeoff, whereas subjects in the delayed condition pressed the “enter” key slower. In the delayed condition, the slowing conclusion key coupled with the faster third keystroke possibly indicates that subjects adopted a chunking strategy at the beginning of training, then switched to a non-chunking strategy in order to expend less energy and processing power.

Testing

Accuracy. There were no significant effects for accuracy during testing, indicating a possible ceiling effect for the test itself. Otherwise, the data do not show that holding information in working memory or forcing subjects to chunk the information improved accuracy.

Overall response latency. A main effect was found for test item type, $F(1, 44) = 6.162$, $MSE = 0.005$, $p = .0169$. Across conditions, typing a new item was slower ($M = 2.478$, $SEM = .082$) than typing an old item ($M = 2.441$, $SEM = .082$), reflecting repetition priming. There was a significant interaction for test item type and training condition, $F(1, 44) = 4.380$, $MSE = 0.005$, $p = .0422$. Subjects in the delayed condition were slower overall and did not show improved memory for the old items ($M = 2.504$, $SEM = .099$) compared to the new items ($M = 2.510$, $SEM = .099$). Subjects in the immediate condition were faster for the old items ($M = 2.379$, $SEM = .131$) than for the new items ($M = 2.446$, $SEM = .132$).

Initiation. A significant interaction was found for training condition and training gap presence, $F(1, 44) = 5.907$, $MSE = 0.481$, $p = .0192$. Subjects in the delayed condition who were presented with a gap in the four-digit number during training typed faster ($M = 1.183$, $SEM = .062$) on the first keystroke during testing than subjects in the delayed condition who were presented without the gap during training ($M = 1.489$, $SEM = .090$), whereas subjects in the immediate condition typed faster on the first keystroke without the gap ($M = 1.214$, $SEM = .085$) than with the gap ($M = 1.596$, $SEM = .139$). The main effects of training condition, training gap presence, and test item type were not significant for the first keystroke.

Execution. A main effect was found for training condition, $F(1, 44) = 5.383$, $MSE = 0.006$, $p = .0250$. Overall, subjects in the delayed condition were slower ($M = .246$, $SEM = .005$) than subjects in the immediate condition ($M = .225$, $SEM = .004$). A main effect was found for keystroke, $F(2, 88) = 63.000$, $MSE = 0.003$, $p < .0001$. The third keystroke was significantly slower ($M = .284$, $SEM = .007$) than the second ($M = .220$, $SEM = .004$) and fourth keystrokes ($M = .202$, $SEM = .003$). A post hoc test revealed significant mean differences between the second and third keystroke ($p < .0001$), the second and fourth keystroke ($p = .0215$), and the

third and fourth keystroke ($p < .0001$). Figure 9 shows a significant interaction for keystroke and training gap presence, $F(2, 88) = 4.603$, $MSE = 0.003$, $p = .0125$. Subjects who were presented the gap during training were slower on the third keystroke ($M = .302$, $SEM = .011$) during testing than subjects who were not presented the gap during training ($M = .266$, $SEM = .006$), suggesting that subjects who saw a space adopted a chunking strategy and transferred it to the test. There was no difference between gap and no gap for the second and fourth keystrokes during testing. A main effect was found for test item type, $F(1, 44) = 21.728$, $MSE = 0.0002161$, $p < .0001$. Overall, new items were typed slower ($M = .240$, $SEM = .005$) than old items ($M = .232$, $SEM = .005$), reflecting repetition priming. Figure 10 shows a significant interaction for keystroke and test item type, $F(2, 88) = 14.719$, $MSE = 0.0001973$, $p < .0001$. For the third keystroke, typing a new item was slower ($M = .295$, $SEM = .010$) than typing an old item ($M = .274$, $SEM = .009$), whereas for the second and fourth keystrokes there was no difference between new and old. The main effect of training gap presence was not found to be significant for testing.

Conclusion. The only significant finding was a main effect for training condition, $F(1, 44) = 8.243$, $MSE = 0.009$, $p = .0063$. Subjects in the delayed condition were slower to press the “enter” key during testing overall ($M = .284$, $SEM = .013$) than subjects in the immediate condition ($M = .227$, $SEM = .004$). Training gap presence and test item type were not significant.

Discussion

The purpose of Experiment 2 was to test whether using working memory during training would help performance on a test, as in Experiment 1. Additionally, Experiment 2 sought to test whether presenting a physical space between numbers, or chunking them, would help performance on a test. As in Experiment 1, retention from training to testing was measured using old and new numbers.

During training, accuracy increased across blocks overall. Further analysis indicated that subjects in the immediate condition showed a slight numerical decline in accuracy across training blocks (from .921 to .898), but subjects in the delayed condition improved substantially in accuracy (from .587 to .732; see Figure 5). The initiation and execution keystrokes became faster across blocks overall, however, further analysis indicated that only the third keystroke improved in response latency out of the execution keystrokes (see Figure 8). The second and fourth keystrokes essentially did not change across blocks.

The accuracy and response latency results for training in the delayed condition do not support the speed-accuracy tradeoff in that accuracy actually increased overall (see Figure 5) at the same time that response latency improved overall (see Figure 7), except for the conclusion keystroke where it actually became slower with each block (see bottom panel, Figure 3). If the speed-accuracy tradeoff had an effect, subjects should have declined in accuracy while improving response latency overall. Even the subjects in the immediate condition, with a less demanding training condition, showed only a small decrease in accuracy. Accuracy in the immediate condition was at the ceiling, as seen by the very low variability, so there was no room for improvement and responses were not variable enough to see a substantial decline. In addition, the fact that the number only stayed on the screen for two seconds may have forced the subjects in the immediate condition to focus better than in Experiment 1 or else they would not see the number before it disappeared. The reason subjects in the delayed condition improved in accuracy and response latency could be that they had more room for improvement and the difficulty of the task kept them engaged. The working memory demands of the delayed condition may have served as an effective cognitive antidote in Experiment 2 (Kole et al., 2008). The working

memory demands of the delayed condition in Experiment 1 were apparently effectively mitigated by the strategy chosen by subjects in that case.

Repetition priming was supported in that, overall, new numbers were typed slower than old numbers, especially in the immediate condition. Broken down by keystroke, old numbers were typed slower than new numbers in execution time (see Figure 10). Specifically, subjects typed the third keystroke slower for new numbers than old numbers, and the second and fourth keystrokes essentially showed no difference for item type. Similar to Experiment 1, the third keystroke could be slower for new numbers because of the extra time taken to encode the unfamiliar numbers than numbers that had been trained. Subjects in the immediate condition typed both new and old numbers faster than subjects in the delayed condition, and the immediate condition essentially showed more improvement in total response latency from new to old numbers, whereas the delayed condition essentially showed no improvement from new to old numbers.

The reason for the improvement from new to old numbers in total time for the immediate condition but not for the delayed condition could be the different strategies used in each training condition. If subjects in the immediate condition adopted a chunking strategy, typing the first and third keystroke slower, then they would have had more time to encode the numbers than subjects in the delayed condition. The immediate condition had the full two seconds to look at each number. Also, if subjects in the delayed condition tried to finish typing the number in working memory before the number on the current screen disappeared, they would not have had as much time to encode the current number and, therefore, retention would be almost non-existent. If subjects in the delayed condition had actually used the intended strategy of holding two four-digit numbers in working memory, then they would show no difference from the immediate

condition for initiation time. However, they were much faster on initiation time during training, suggesting the strategy of typing the previous number before the current number disappeared.

As in Experiment 1, contextual interference was not supported. During training, the immediate condition was more accurate and faster overall than the delayed condition in conclusion time (although not in initiation time), which supports contextual interference in that the more difficult condition will perform worse during training. During testing, however, there was no difference between training conditions in accuracy, and the delayed condition typed slower on execution and conclusion time. If contextual interference had an effect, delayed should have performed worse during training in accuracy and response latency, but surpassed the immediate condition in accuracy and response latency during testing. The memory requirement of the delayed condition could have been too difficult to promote any benefit from training to testing. The subjects in their respective training conditions could also have kept the strategy used during training to perform the test: the immediate condition chunking the response and the delayed condition seeing the number as one big chunk.

The hypothesis that using working memory during training would improve performance on a test did not have strong evidence. There was no difference between training conditions for accuracy on the test. On the analysis measuring retention of the specific trained numbers, the immediate condition showed improvement for old numbers compared to new numbers in total response latency, whereas the delayed condition did not show any difference in total time between old numbers and new numbers. The item type result shows no retention of trained numbers in the delayed condition and, therefore, no benefit to using working memory during training. In fact, the results show that using working memory during training could potentially be harmful to future performance.

Each training condition exhibited different strategies similar to Experiment 1. The immediate condition took longer on the first keystroke than the delayed condition during training. In execution, subjects typed the third keystroke slower than the second and fourth keystrokes during training and testing. The immediate condition typed the conclusion faster than the delayed condition for training and testing. In addition, the delayed condition became faster on the third keystroke across blocks (see Figure 8) and the conclusion for the delayed condition became slower across blocks during training (see bottom panel, Figure 3). These results suggest that the immediate condition adopted a chunking strategy in which the first and third keystrokes contained the time to encode each four-digit number. The immediate condition also probably rushed into the conclusion keystroke, meaning subjects chunked the five keystrokes into the first two numbers and the last two numbers plus the “enter” key.

The delayed condition likely tried to finish typing the number in working memory before the current number disappeared, as evidenced by the fast rather than slow initiation time. However, there was no difference between the training conditions in execution time, meaning the delayed condition took a longer time on the third keystroke. A possible strategy that the subjects in the delayed condition adopted could be to type the first two numbers immediately when the next screen appeared, then look at the number on the current screen, then type the last two numbers and the “enter” key. This strategy would explain the fast initiation and slow third key, but not the increase in speed of the third keystroke across blocks and the decrease in speed of the “enter” key across blocks. Perhaps the strategy changed as subjects in the delayed condition trained. Maybe they began training with the suggested strategy stated above: typing the first two numbers immediately when the next screen appeared, looking at the new number, and typing the last two numbers and pressing the “enter” key. As training progressed and they became more

proficient with using the 10-keypad, they could have realized that the chunking strategy made the task more difficult than typing all four numbers very quickly immediately on the new screen, looking at the new number, and then pressing the “enter” key. The change in strategy for the delayed condition would explain the results above as well as the third keystroke decreasing in response latency and the conclusion keystroke increasing in response latency across blocks. They would have switched from a chunking strategy similar to the immediate condition to a strategy that saw the four-digit number as one big chunk.

Overall, inserting a physical space between the second and third numbers caused the subjects who were presented with the space to have slower execution times than subjects who were not presented with the space during training. Specifically, there was essentially no difference between having the space and not having the space for the second and fourth keystrokes, but there was a large difference for the third keystroke during training (see Figure 6). Subjects who saw the space were slower on the third keystroke, suggesting that presenting the number with a space enhanced the chunking effect. However, the hypothesis that forcing subjects to chunk during training would enhance performance on a test did not prove true. There was no difference between training with the space than without the space for accuracy on the test. For initiation time (and total response time), training with the space helped test performance in the delayed condition but hurt test performance in the immediate condition. During testing, subjects who saw a space during training typed the third keystroke slower than subjects not presented with a space, similar to the training result (see Figure 9). The fact that the slower third keystroke for subjects presented with a space transferred to testing indicates that they successfully adopted the chunking strategy, but it did not help them perform any better on the test overall.

The unexpected result from testing is that subjects in the delayed condition who saw numbers with a space typed faster on the first keystroke (and faster in total response latency) than subjects not presented with a space in the numbers. In contrast, subjects in the immediate condition who saw numbers with a space were slower on the first keystroke (and slower in total response latency) than subjects in the same condition not presented with a space. If subjects in the immediate condition used a chunking strategy without seeing a space, then seeing a space may have encouraged more time to encode during not just the third keystroke but also the first keystroke. The reverse result for the delayed condition is more difficult to find an explanation. A tentative explanation could be that subjects in the delayed condition who saw a space between the numbers during training successfully learned to chunk, not just in their typed response but also in memory. If they looked at the first two numbers then started typing, they would not need as long to encode because they encoded two numbers instead of four. In sum, presenting numbers with a space in the middle encouraged subjects to respond by typing numbers with more time between them, but it may not have encouraged subjects to mentally memorize the numbers as chunks and, therefore, did not increase retention of the trained numbers. Presenting gaps in stimuli simply provokes a similar response from subjects, as mentioned by O'Shea and Clegg (2006).

Experiment 2 had some flaws that may have affected the outcome. Limiting the amount of time that the numbers stayed on the screen may not have induced the usage of working memory to the intended extent, but simply changed the strategy used. As such, the data entry task may need to be modified in a different way to guarantee that subjects actually use working memory as intended. Presenting a space between the numbers may not have induced memorizing the numbers as chunks, so a different strategy may need to be used in order to force subjects to

chunk the numbers in memory and corroborate the findings in this study that chunking did not help performance overall. Even if subjects did memorize the numbers in chunks as intended, the limited time that the numbers appeared on the screen may have hindered subjects in the delayed condition from being able to fully encode the chunks because they had less time to look at the new number than subjects in the immediate condition.

General Discussion

Two experiments sought to test whether training while using working memory would improve performance on a test using a data entry task. Both experiments used new and old numbers in the test to measure retention from training to testing. Experiment 1 revealed evidence to support the speed-accuracy tradeoff in that subjects overall became less accurate but faster in speed across blocks in training. Experiment 2 found some evidence for the speed-accuracy tradeoff in that the immediate condition slightly declined in accuracy across blocks, and decreased response latency overall throughout each block in training. However, the delayed condition increased in accuracy (see Figure 5) and decreased response latency (see Figure 7) across blocks, except for the conclusion keystroke that became slower, suggesting that the speed-accuracy tradeoff had no overall effect for the delayed condition because the task was more difficult by reason of the time limit that the numbers stayed on the screen.

Kole et al. (2008) suggested that implementing a cognitive antidote to inhibition that affects accuracy (boredom, mental fatigue, and physical fatigue) could keep accuracy from decreasing over time. One antidote, as seen in the current Experiment 2, could be to use cognitive complications to prevent inhibition. In the context of Experiment 2, the delayed condition increased in accuracy over time because the subjects' cognitive abilities were taxed with trying to hold two numbers in working memory, which kept them focused and engaged on

the task, or served as an antidote to their inhibition. Overall, the experiments support the speed-accuracy tradeoff, but only if the task is easy enough that the subjects start to disengage their attention from the task.

Repetition priming was partially supported in that subjects typed new numbers slower than old numbers for the third keystroke during the test in both experiments. The third keystroke contains an extra encoding step in the data entry task, while the first keystroke contains encoding and motor planning (Fendrich et al., 1991). If repetition priming had an effect, subjects would need more time to encode the new numbers than the old numbers, as seen by the results for the third keystroke. In Experiment 2, for total response time the immediate condition typed old numbers faster than the delayed condition, and the delayed condition showed essentially no difference between typing old and new numbers. The difference between the training conditions for item type in Experiment 2 could be the different strategies used during training and the fact that subjects in the delayed condition did not have as much time to encode each number because they typed the previous number as they tried to encode the current number before it disappeared. Subjects in the immediate condition adopted a chunking strategy and, therefore, had more time to encode during the first and third keystroke. More time to encode translates to better memory for numbers trained during the test.

Extending the concept of contextual interference to include the delayed condition in the current study does not present evidence to support the concept. During training in Experiment 1, the delayed condition improved more in total response latency across blocks of training because subjects had more room for improvement, but the delayed condition was still numerically slower than the immediate condition overall (see Figure 1). In Experiment 2, the delayed condition was less accurate during training (see Figure 5). Slower response latency and less accuracy agree

with contextual interference in that the more difficult condition, the condition that ideally would have more interference with training a task, led to worse performance during training. However, there was no difference in accuracy between the training conditions at test for both experiments. In fact, the delayed condition still typed slower for execution and conclusion at test in Experiment 2, even though both training conditions completed the same test. Experiment 2 found evidence that the immediate condition also had better retention because the immediate condition typed faster in total response latency for old numbers than the delayed condition, and the immediate condition showed more improvement from new numbers to old numbers. The concept of contextual interference did not find support in the current study, possibly because in Experiment 1 both training conditions did not have room to improve in accuracy, and in Experiment 2 the delayed condition was too difficult for any retention to occur.

Both experiments demonstrated possible strategies that each condition used to progress through training and testing. During training, both experiments revealed that the immediate condition adopted a chunking strategy, where subjects took more time to type the first and third keystroke in order to encode the numbers, and the encoding showed in their chunked responses. Specifically, the immediate condition chunked the five total keystrokes into the first two and then the last three, taking more time on the first and third and sliding into the “enter” key. When the space between numbers was introduced in Experiment 2, the chunking became even more pronounced in the immediate condition for the third keystroke.

The delayed condition in Experiment 1 seemed to adopt a strategy where subjects looked at the four-digit number as one big chunk and did not take any more time to encode during the third keystroke. The encoding was done in the conclusion keystroke because subjects immediately typed the number in working memory as soon as the new screen appeared, then

looked at the current number on the screen before pressing “enter.” In Experiment 2, the delayed condition began with a chunking strategy as seen by the slower third keystroke, then seemed to switch to the strategy of seeing the four-digit number as one big chunk used in Experiment 1 because the third keystroke became faster (see Figure 8) and the conclusion became slower (see bottom panel, Figure 3) across blocks. Once the act of maneuvering around the 10-keypad became easier, subjects in the delayed condition switched strategies so that they could type the number in working memory immediately on the new screen and look at the current number before it disappeared in an attempt to reduce the amount of effort involved with using working memory.

The major hypothesis of the current study was to see whether using working memory while training would lead to better performance on a test. Both experiments did not find a difference in accuracy between training conditions at test, possibly because the test was designed in such a way that neither training condition could get any more accurate. Experiment 2 found some evidence that training while using working memory may actually hinder performance because the delayed condition typed slower for execution and conclusion time than the immediate condition at test. The immediate condition also typed old numbers faster overall (i.e., for total response latency) than the delayed condition in Experiment 2, suggesting the immediate condition had better retention because of the extra encoding time available to the immediate condition during training. Overall, training while using working memory did not help performance on the test, but further modifications to the task need to be undertaken in order to definitively reject the hypothesis.

Experiment 2 introduced a space in the four-digit numbers in order to test whether forcing subjects to chunk would lead to better performance on the test. Presenting a space in the

numbers only increased the response latency for the third keystroke (see Figures 6 and 9), as if subjects imitated what they saw (O'Shea & Clegg, 2006). There was no difference in accuracy at test between training with a space and training without a space. The puzzling result was that the delayed condition typed faster on initiation with a space, whereas the immediate condition typed slower on initiation with a space during the test. The explanation given was that the immediate condition was using the space to signal to the subjects to take more time to encode on the first and third keystroke, whereas subjects in the delayed condition presented with a space successfully learned to chunk. They could have encoded only the first two numbers, which would explain the faster initiation, and then encoded the last two numbers, instead of encoding all four numbers first as in the immediate condition.

The current study has posed a novel question and found few results answering that question. Some future experiments could change the data entry task so that subjects have to use their working memory and also forced to chunk. One modification could be to have subjects memorize two screens of numbers before they type the number from the first screen. This modification would force subjects to use their working memory in the intended manner, holding two numbers in working memory while typing one and memorizing another. Another modification could be to make the numbers longer so that working memory is more taxed. Longer numbers could also force subjects to think of the number as smaller chunks, which should help with memorizing the number and performing better on the test. A modification to ensure chunking could be to present numbers on separate screens in the desired chunks so that subjects see the numbers separately, as chunks, and hopefully memorize them as chunks. Then they would be asked to type the separate chunks at once. A different measure that could be utilized in order to see exactly the strategy that subjects adopt could be to track accuracy for each

keystroke rather than counting a whole four-digit number as wrong if subjects miss one keystroke. Finally, designing a harder test would be ideal, for example presenting the numbers as words or making the numbers longer, so that subjects in all training conditions can actually show improvement and to be able to see a difference between the training conditions at test.

References

- Bourne, L. E., Jr., Healy, A. F., Bonk, W. J., & Buck-Gengler, C. J. (2011). Intention to respond in a special way offers some protection against forgetting associations. *American Journal of Psychology*, 124, 23-36. doi: 10.5406/amerjpsyc.124.1.0023
- 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. doi: 10.1037/0278-7393.27.3.879
- Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin and Review*, 17, 193-199. doi: 10.3758/PBR.17.2.193
- Chooi, W., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, 40, 531-542. doi: <http://0-dx.doi.org/libraries.colorado.edu/10.1016/j.intell.2012.07.004>
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-183. doi: <http://0-dx.doi.org/libraries.colorado.edu/10.1017/S0140525X01003922>
- Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science*, 208, 1181-1182. doi: <http://0-dx.doi.org/libraries.colorado.edu/10.1126/science.7375930>
- 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. doi: 10.1037/0278-7393.17.1.137

- 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. doi: 10.1037/1076-898X.10.3.188
- Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne, L. E., Jr. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 534-546. doi: 10.1037/0278-7393.32.3.534
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). *Proceedings of the National Academy of Sciences of the United States of America*, 105, 6829-6833. doi: 10.1073/pnas.0801268105
- Kole, J. A., Healy, A. F., & Bourne, L. E., Jr. (2008). Cognitive complications moderate the speed-accuracy tradeoff in data entry: A cognitive antidote to inhibition. *Applied Cognitive Psychology*, 22, 917-937. doi: 10.1002/acp.1401
- Kole, J. A., Healy, A. F., & Buck-Gengler, C. J. (2005). Does number data entry rely on the phonological loop? *Memory*, 13, 388-394. doi: 10.1080/09658210344000224
- Kundu, B., Sutterer, D. W., Emrich, S. M., & Postle, B. R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *The Journal of Neuroscience*, 33, 8705-8715. doi: <http://0-dx.doi.org/libraries.colorado.edu/10.1523/JNEUROSCI.5565-12.2013>
- Lee, J. L. C. (2008). Memory reconsolidation mediates the strengthening of memories by additional learning. *Nature Neuroscience*, 11, 1264-1266. doi: 10.1038/nn.2205
- Lee, J. L. C. (2009). Reconsolidation: Maintaining memory relevance. *Trends in Neuroscience*, 32, 413-420. doi: 10.1016/j.tins.2009.05.002
- Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill

- acquisition. *Journal of Experimental Psychology*, 9, 730-746. doi:
<http://0-dx.doi.org.libraries.colorado.edu/10.1037/0278-7393.9.4.730>
- Lohse, K. R., & Healy, A. F. (2012). Exploring the contributions of declarative and procedural information to training: A test of the procedural reinstatement principle. *Journal of Applied Research in Memory and Cognition*, 1, 65-72. doi: 10.1016/j.jarmac.2012.02.002
- Machizawa, M. G., & Driver, J. (2011). Principal component analysis of behavioural individual differences suggests that particular aspects of visual working memory may relate to specific aspects of attention. *Neuropsychologia*, 49, 1518-1526. doi:
10.1016/j.neuropsychologia.2010.11.032
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97. doi:
10.1037/h0043158
- O'Shea, G., & Clegg, B. A. (2006). Stimulus and response chunking in the Hebb digits task. *Psychological Research*, 70, 180-192. doi: 10.1007/s00426-004-0198-1
- Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, 20, 713-721. doi: 10.1002/acp.1224
- Ridgeway, D. (2006). Strategic grouping in the spatial span memory task. *Memory*, 14, 990-1000. doi: 10.1080/09658210601010797
- Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology*, 5, 179-187. doi:
<http://0-dx.doi.org.libraries.colorado.edu/10.1037/0278-7393.5.2.179>

von Bastian, C. C., Langer, N., Jäncke, L., & Oberauer, K. (2013). Effects of working memory training in young and old adults. *Memory and Cognition*, 41, 611-624. doi:

10.3758/s13421-012-0280-7

Wohldmann, E. L., Healy, A. F., & Bourne, L. E., Jr. (2008). Global inhibition and midcourse corrections in speeded aiming. *Memory & Cognition*, 36, 1228-1235. doi:

10.3758/MC.36.7.1228

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, 9, 185-211. doi:

10.3758/BF03196276

Young, M. D., Healy, A. F., Gonzalez, C., Dutt, V., & Bourne, L. E., Jr. (2011). Effects of training with added difficulties on RADAR detection. *Applied Cognitive Psychology*, 25,

395-407. doi: 10.1002/acp.1706

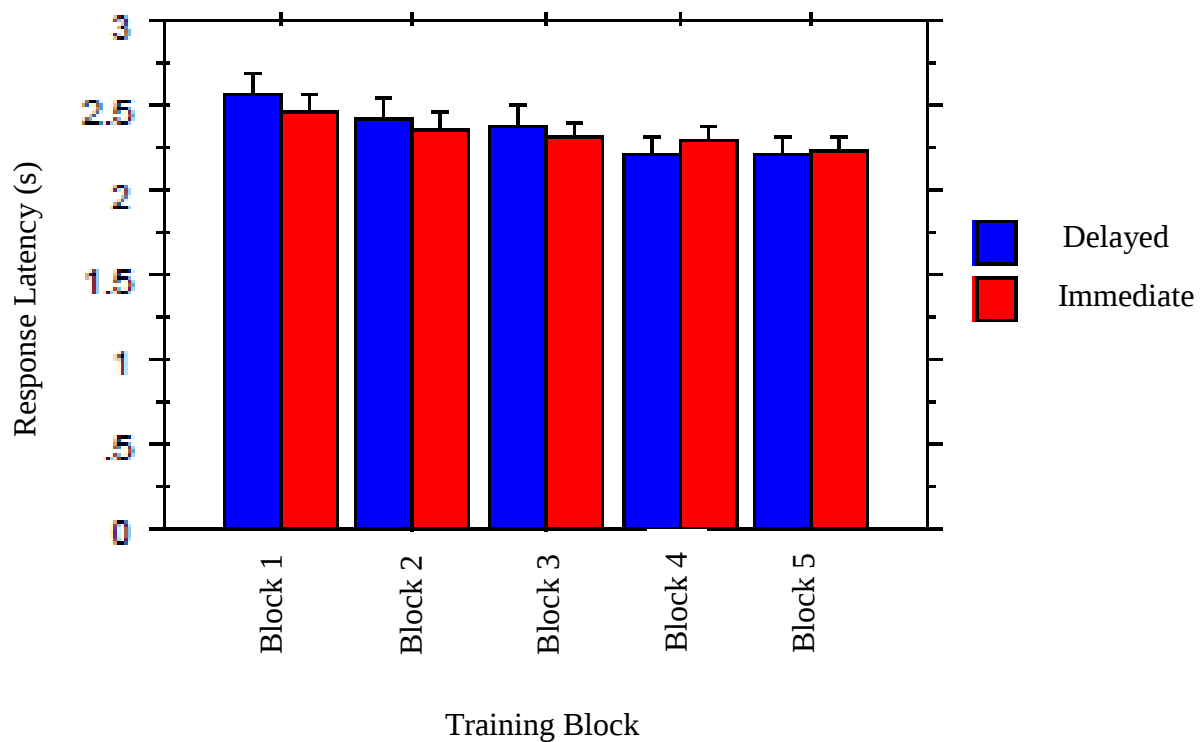


Figure 1. Total response latency for training by training block and training condition, Experiment 1. Bars represent standard errors of the mean.

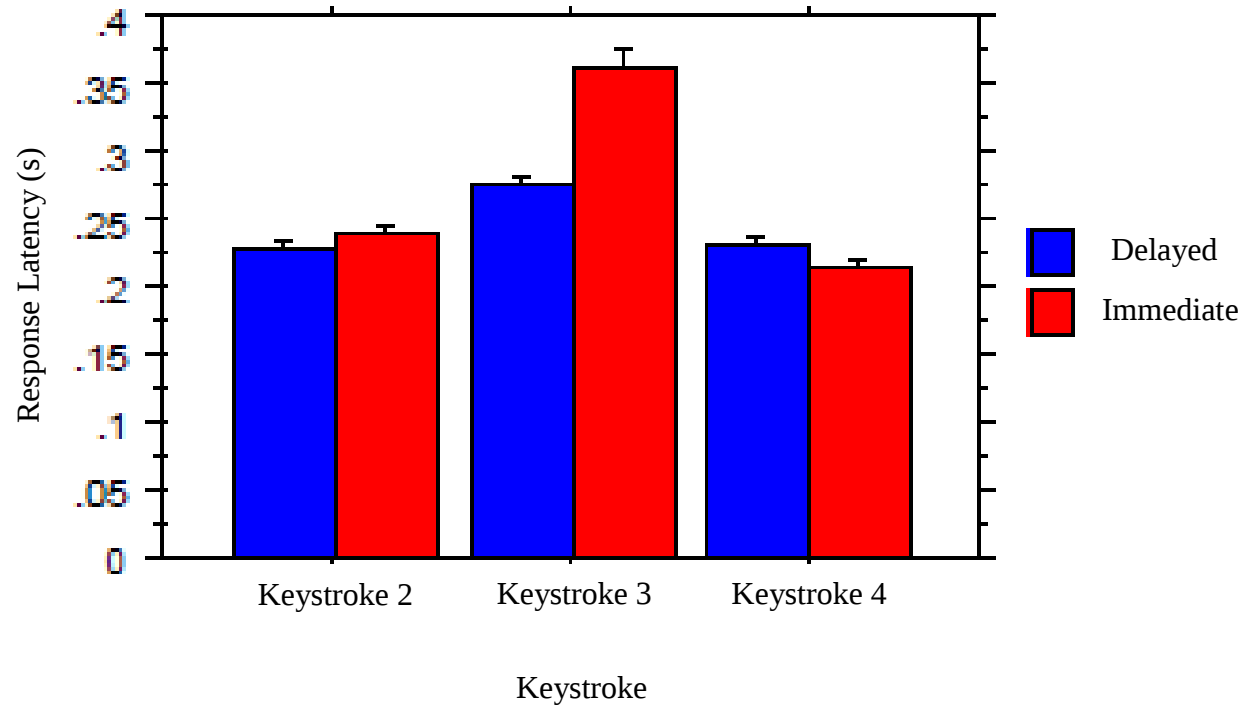


Figure 2. Execution time for training by keystroke and training condition, Experiment 1. Bars represent standard errors of the mean.

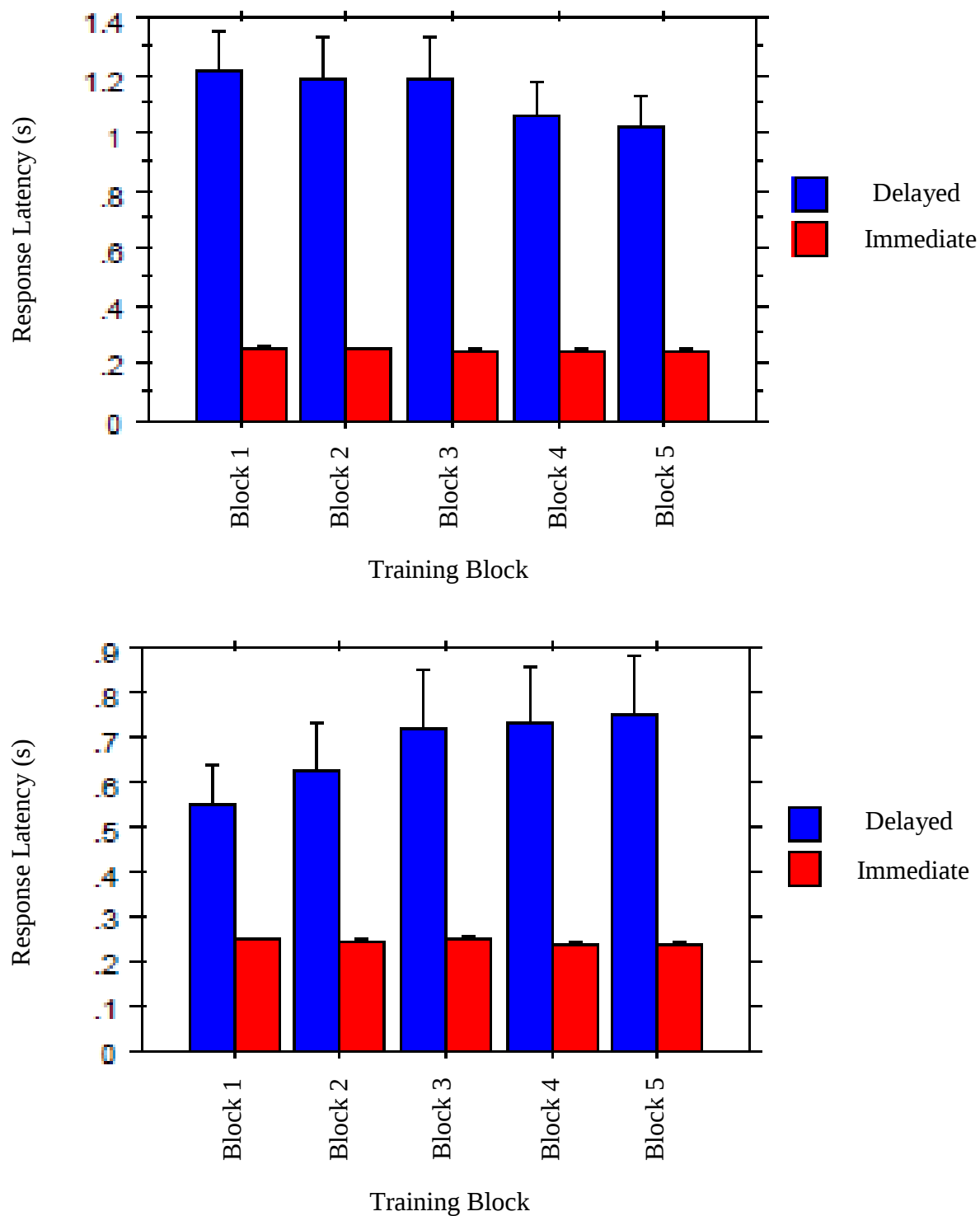


Figure 3. Conclusion time for training by training block and training condition for Experiment 1 (top panel) and Experiment 2 (bottom panel). Bars represent standard errors of the mean.

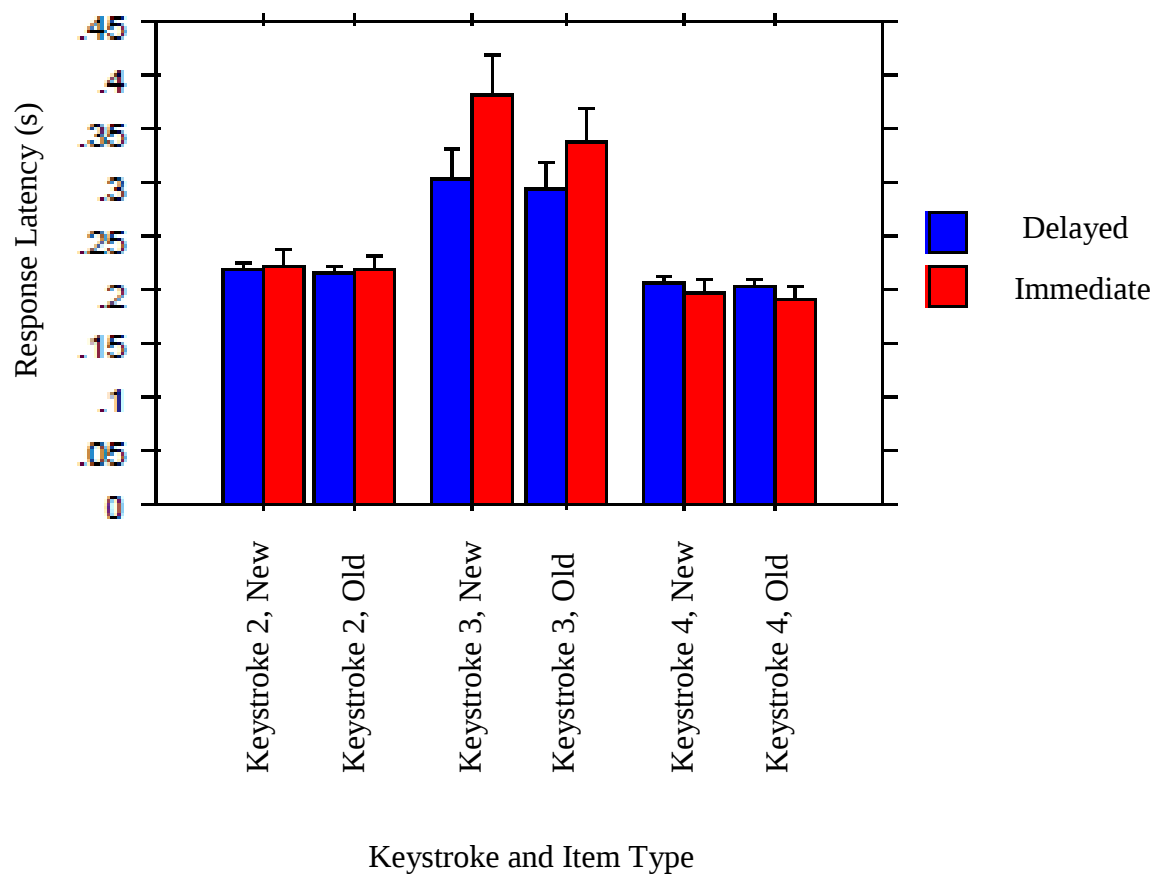


Figure 4. Execution time for testing by keystroke, item type, and training condition, Experiment 1. Bars represent standard errors of the mean.

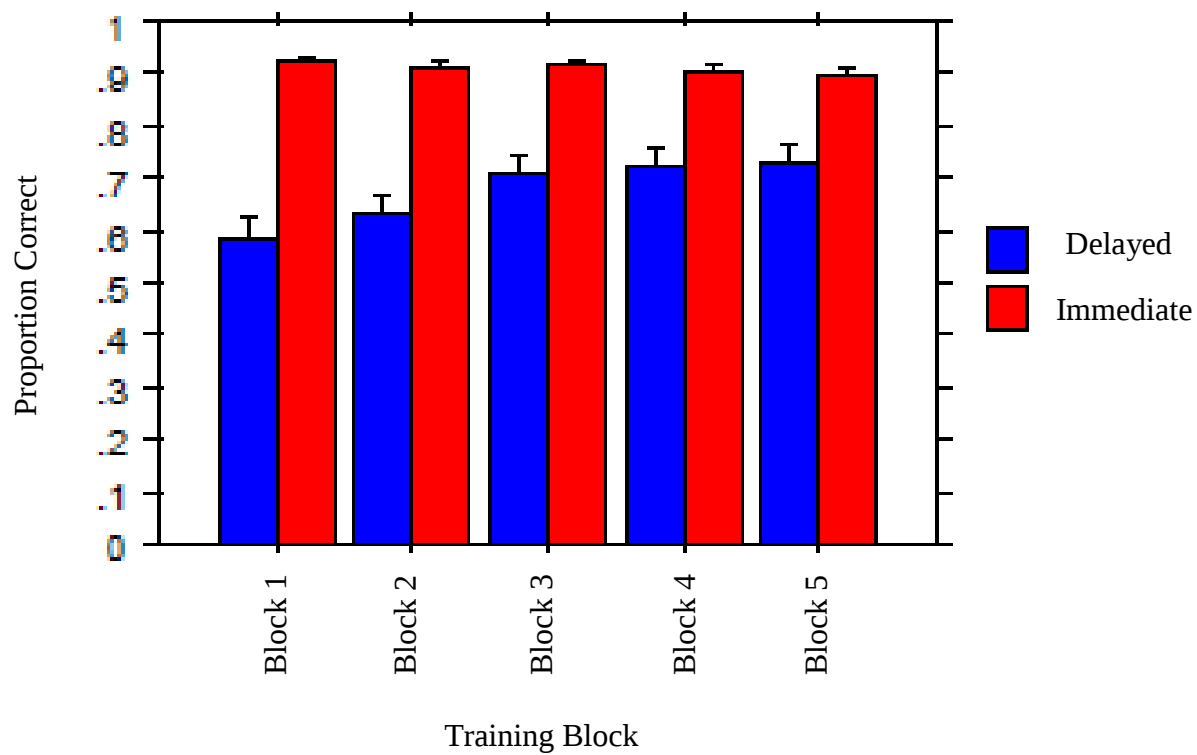


Figure 5. Total accuracy for training by training block and training condition, Experiment 2. Bars represent standard errors of the mean.

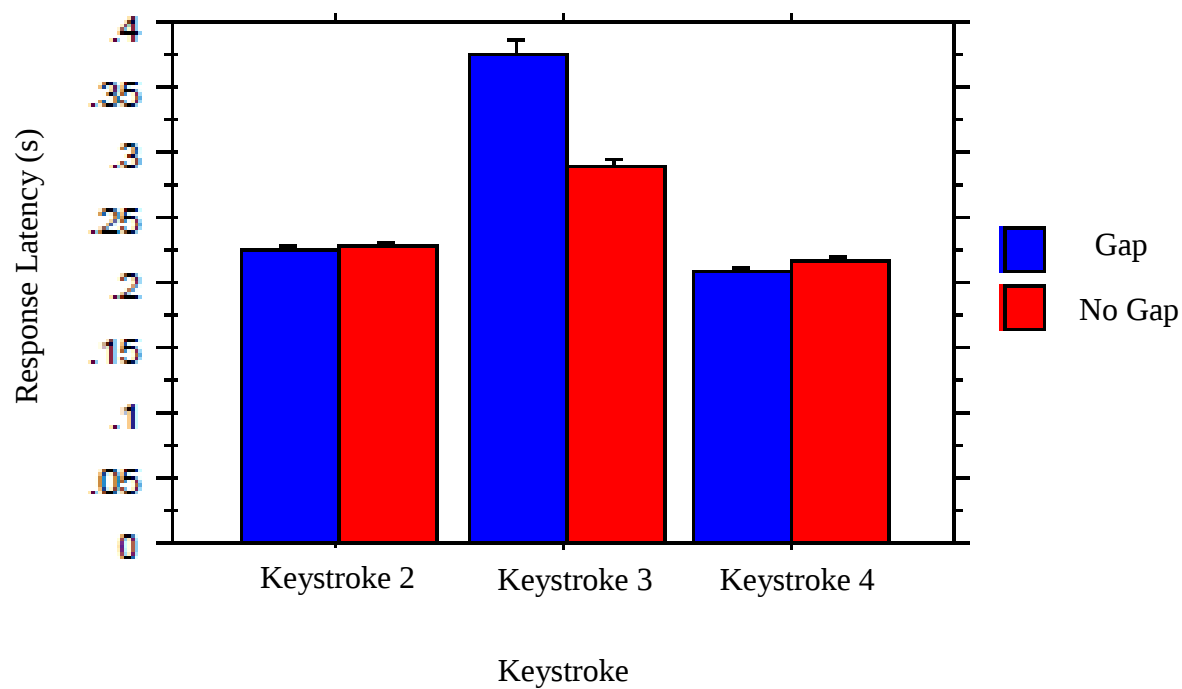


Figure 6. Execution time for training by keystroke and training gap presence, Experiment 2. Bars represent standard errors of the mean.

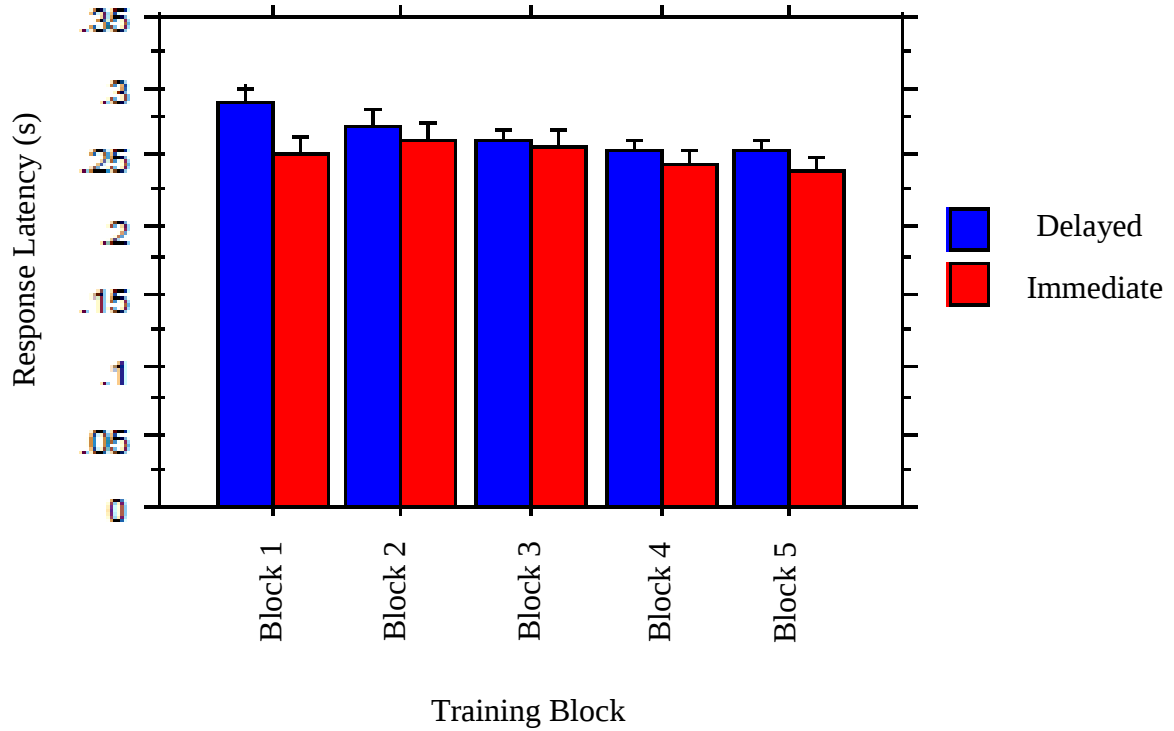


Figure 7. Execution time for training by training block and training condition, Experiment 2. Bars represent standard errors of the mean.

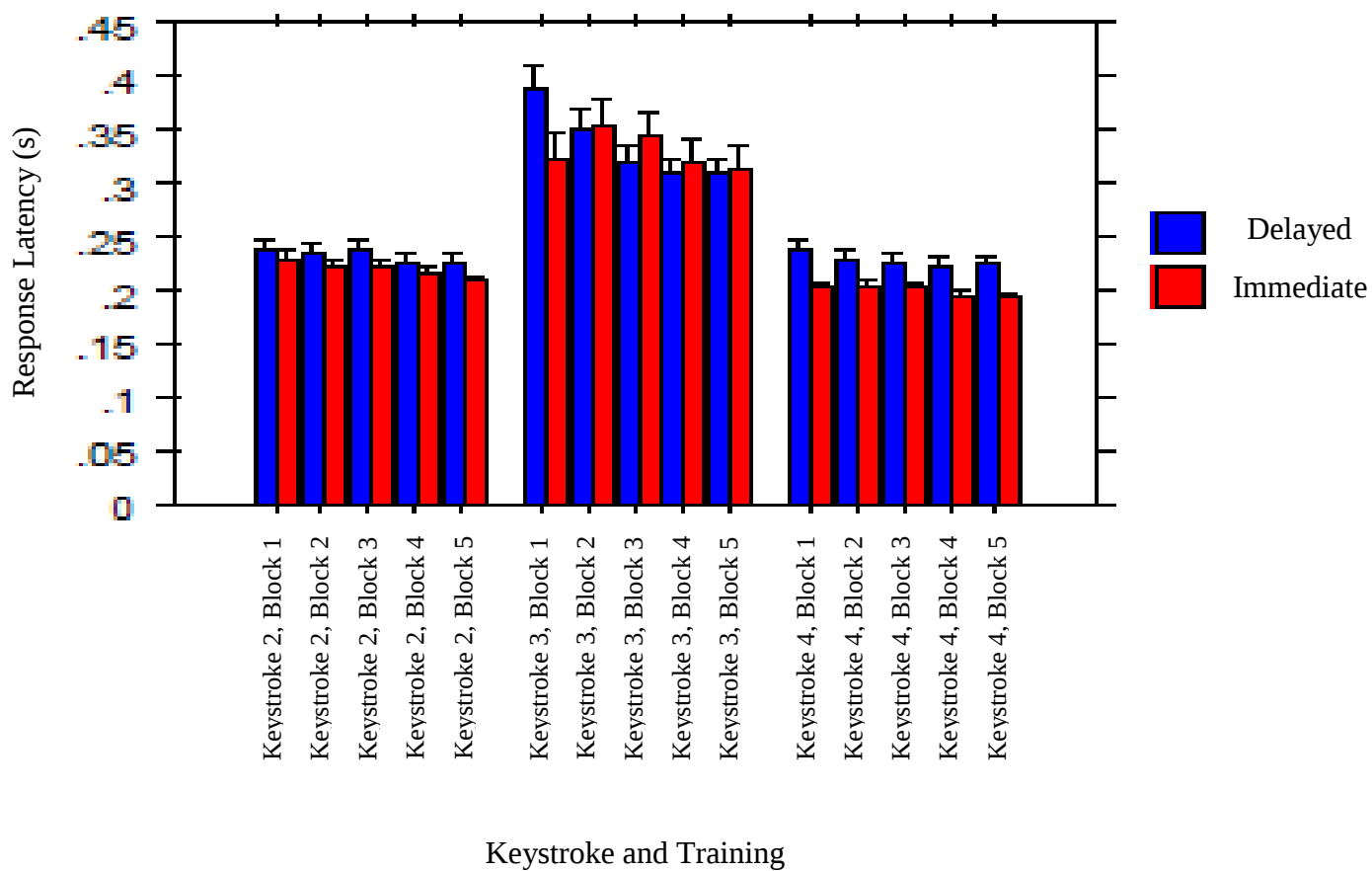


Figure 8. Execution time for training by keystroke, training block, and training condition, Experiment 2. Bars represent standard errors of the mean.

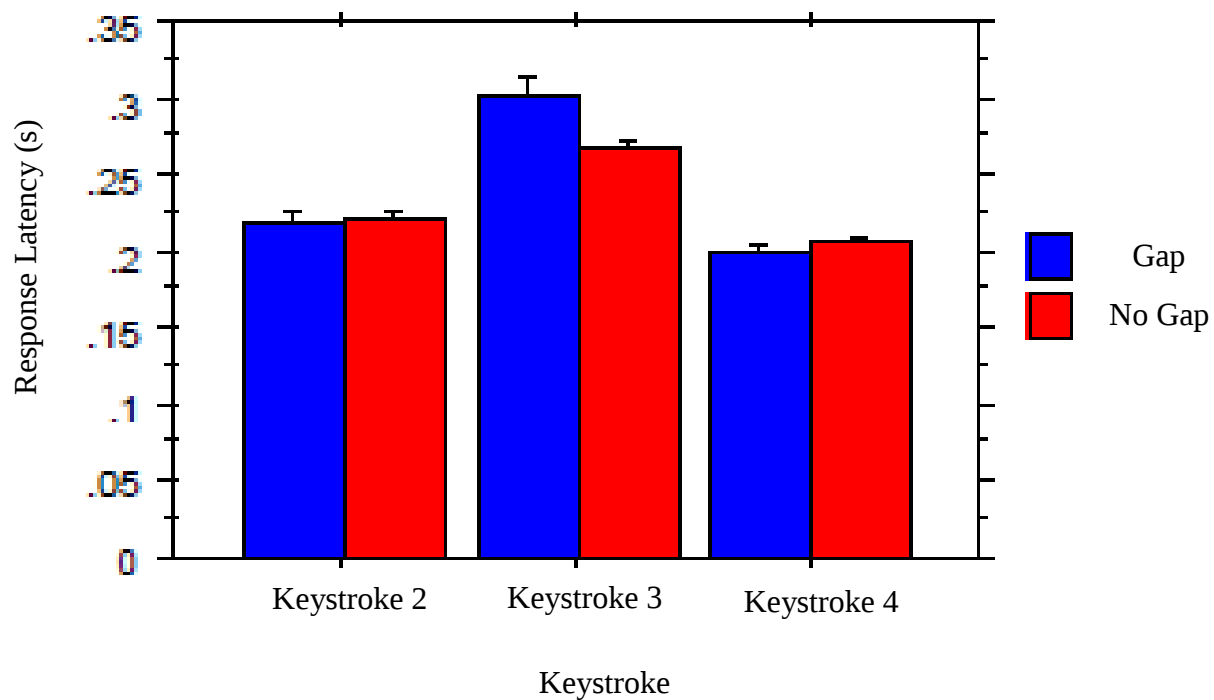


Figure 9. Execution time for testing by keystroke and training gap presence, Experiment 2. Bars represent standard errors of the mean.

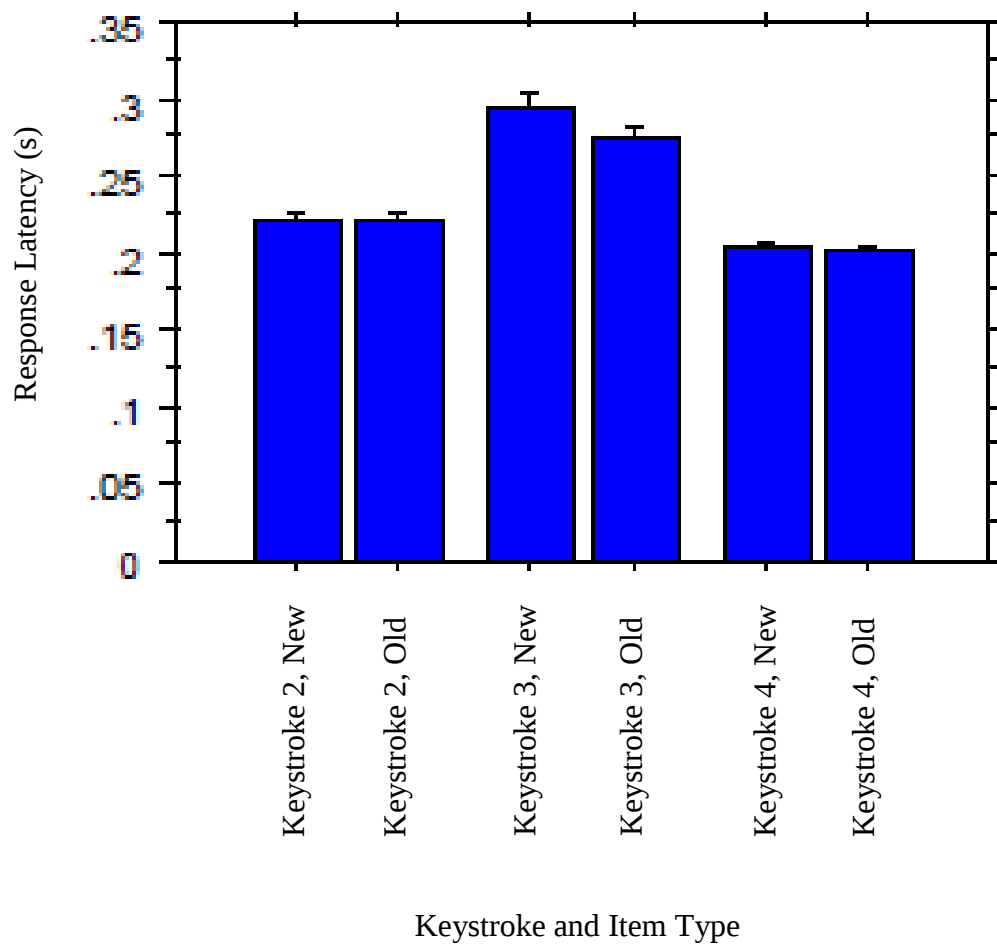


Figure 10. Execution time for testing by keystroke and item type, Experiment 2. Bars represent standard errors of the mean.