SAND2014-17530J

Practice Makes Imperfect: Working Memory Training Can Harm Recognition Memory Performance

Laura E. Matzen, Michael C. Trumbo, Michael J. Haass, Austin Silva, Susan M. Stevens-Adams

Sandia National Laboratories

Michael F. Bunting, Polly O'Rourke

Center for the Advanced Study of Language

Corresponding author:

Laura E. Matzen

Cognitive Systems Research and Applications

Sandia National Laboratories

P.O. Box 5800

Albuquerque, NM87185-1011

Phone: (505) 844-1505

E-mail: lematze@sandia.gov

Abstract

There is a great deal of debate concerning the benefits of working memory (WM) training and whether or not that training can transfer to other tasks. While a consistent finding is that WM training programs elicit a short-term near transfer effect (i.e., improvement in WM skills), results are inconsistent when considering persistence of such improvement and far transfer effects. In the present study, we compared three groups of participants: a group that received WM training, a group that received training on how to use a mental imagery memory strategy, and a control group that received no training. While the WM training group improved on the trained task, their post-training performance on non-trained WM tasks did not differ from that of the other two groups. In addition, while the imagery training group's performance on the recognition memory tasks increased after training, the WM training group's performance on the recognition memory tasks *decreased* after training. Participants' descriptions of the strategies that they used to remember the studied items indicated that WM training may lead people to adopt memory strategies that are less effective for other types of memory tasks. These results indicate that WM training may have unintended consequences for other types of memory performance.

Keywords: working memory training; mental imagery; recognition memory; memory strategies

Practice Makes Imperfect: Working Memory Training Can Harm Recognition Memory Performance

Working memory (WM) refers to the brain system used for storage and manipulation of transitory information necessary for complex tasks such as learning, reasoning, and language comprehension (Becker & Morris, 1999). Recent research has indicated that WM training, where people repeatedly practice increasingly difficult WM tasks, can improve both WM capacity and other aspects of cognitive performance (cf. Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). Improved performance on a trained task should bolster performance for additional domains and tasks to the extent that they rely on overlapping cognitive abilities or share neural systems (Dahlin et al., 2008). As WM has been identified as a central component of general cognition (Engle, Tuholski, Laughlin, & Conway, 1999), the conjecture is that improvement on the trained WM task will result not only in near transfer (improvement on other WM tasks, such as a spatial WM task following training of a verbal WM task, indicative of heightened WM capacity), but also in far transfer (improvements in other domains, such as fluid intelligence tests).

Fluid intelligence refers to those aspects of intelligence that allow for adaptive reasoning and problem solving (Carpenter, Just, & Shell, 1990), and is a construct thought to strongly relate to academic achievement (Rohde & Thompson, 2007). Previous research has demonstrated that WM capacity and fluid intelligence are strongly related constructs, sharing approximately 50% of their variance (Kane, Hambrick, & Conway, 2005). Fluid intelligence was largely thought to be immutable, but a study by Jaeggi and colleagues (2008) showed that WM training could improve fluid intelligence performance. This finding is significant from a theoretical as well as a practical perspective, and many additional studies have investigated the benefits of WM training for fluid intelligence and other types of cognitive processes. These studies have shown that WM training can improve episodic memory (Rudebeck et al., 2012), attention (Chein & Morrison, 2010), and can provide general cognitive enhancement for children (Alloway, Bibile, & Lau, 2013). Additionally, WM training has been proposed as a remediating intervention for populations such as adults with amnestic mild cognitive impairment (Carretti, Borella, Fostinelli, & Zavagnin) and children with dyslexia (Luo et al., 2013) or ADHD (Holmes et al., 2010).

Given the high hopes for improving cognitive abilities through WM training, there has been a proliferation of commercial cognitive training programs (e.g., Brain Age, CogMed, Lumosity, Mindsparke Brain Fitness Pro, Posit Science Brain Fitness, Posit, WMPro) that are largely based on adaptive WM tasks (Melby-Lervåg & Hulme, 2013). Such programs have become a multi-million dollar industry (Aamodt & Wang, 2010), with programs such as CogMed (http://www.cogmed.com) available in over 30 countries and widely used in schools and clinical settings, and Lumosity (http://www.lumosity.com) boasting upwards of 40 million registered users.

However, other research paints a less optimistic view of the benefits of WM training. The idea that WM training can improve fluid intelligence runs contrary to more than a century of research on cognitive training within psychological and educational science. Numerous studies have demonstrated that although task-specific performance commonly increases with training, transfer of this learning to other tasks or domains is rare (Chase & Ericsson, 1981; Ericsson & Delaney, 1998; Healy, Woholdmann, Sutton, & Bourne, 2006; Singley & Anderson, 1989; Thorndike & Woodworth, 1901). Several recent studies of WM training have failed to find near or far task transfer (Chooi & Thompson, 2012; Redick et al., 2013; Thompson et al., 2013) and a recent meta-analysis concluded that while WM training consistently produces near transfer effects, these effects tend to be short-lived and improvement fails to generalize to other domains (Melby-Lervåg & Hulme, 2013). Furthermore, researchers have cited a number of methodological concerns within the WM training field, including use of single tasks to define WM change, inconsistent use of valid WM tasks, comparison of trained groups to a no-contact control group, and subjective measurements of change (Shipstead, Redick, & Engle, 2012).

Given the mixed results regarding the transfer of WM training to other tasks and the ability of WM training to improve fluid intelligence, it seems prudent to approach this topic with a dose of skepticism. In the present study, we sought to examine the effects of WM training on both untrained WM tasks and on a verbal recognition memory task. While WM training has been touted as improving academic success, there has been little research on how WM training impacts other types of memory that are also crucial for learning, such as recognition memory. In one of the few studies to test the effects of WM training on recognition memory, Rudebeck and colleagues (2012) showed that spatial WM training improved performance on visual recognition memory tasks. However, it remains unclear whether or not WM training can improve verbal recognition memory, a type of memory that is particularly important in most educational settings.

With this study, we also sought to address some of the methodological issues that have made it difficult to interpret the results of prior WM training studies (cf. Redick et al., 2013; Shipstead et al., 2010). Specifically, this study used both a no-contact control group and an active control group in which participants were trained to use mental imagery as a memory strategy. Participants were assigned semi-randomly to the three experimental groups (the groups were balanced by gender and age). We selected mental imagery training as the active control condition because mental imagery is long-established as an effective technique for improving recognition memory performance (Pavio, 1971; Prestianni & Zacks, 1974). In theory, both training techniques should improve recognition memory performance, but it was unclear how they would stack up against one another in practice. The two types of training are fundamentally different in terms of the role of memory strategies in the training tasks. Using mental imagery as a memory aid is clearly a strategy, whereas there is debate about the role of strategy in WM tasks. While several studies have found that strategy use improves performance on WM tasks (McNamara & Scott, 2001; Turley-Ames & Whitfield, 2003), the fundamental goal of WM training is to enhance a basic cognitive ability that can translate to improved performance in tasks that were not trained. Thus, adaptive WM training regimens intentionally discourage participants from developing task-specific memory strategies (e.g. Jaeggi et al., 2008).

Task Selection and Hypotheses

All participants completed the same battery of memory tasks before and after training. The battery included a verbal WM task (listening span), a spatial WM task (rotation span), and a verbal recognition memory task. The listening span and rotation span tasks come from the tradition of complex span tasks (CSTs) which are well-established measures of WM capacity (Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005; Daneman & Carpenter, 1980). Relative to "simple" span tasks that require participants to recall a list of stimuli following a brief retention interval, CSTs require the performance of an additional and often unrelated task (e.g., evaluating mathematical equations) concurrently with the main memory task. The addition of this ancillary task means that successful performance on a CST requires both short-term storage and processing of information, meeting the basic definition of WM as simultaneous storage and processing (Baddeley, 2003).

The recognition memory task was designed to include several conditions with varying levels of difficulty. There were words that were studied only once, words that were repeated at short and long lags, and words that were quizzed within the study block. We expected that the testing effect (Karpicke & Roediger, 2008) would make the quizzed words easiest to remember on the subsequent memory test, while the words that were repeated but not quizzed would be more difficult to remember. For the repeated words, we expected that the spacing effect (Melton, 1970) would lead to better performance for the words that were repeated after a longer lag. We expected the poorest memory performance for the words that were studied only once.

The study participants were semi-randomly assigned to one of three memory training groups. The participants in the control group received no memory training. The participants placed in the WM training group completed a series of training sessions that consisted of an adaptive N-back task and an adaptive symmetry span task. These tasks are similar to those that have been used in prior WM training studies (cf. Melby-Lervåg & Hulme, 2013), and were intended to target both verbal and spatial WM. In accordance with past research and with commercial memory training programs, the tasks adapted their difficulty based on the participant's performance.

The participants in the mental imagery training group were trained to create vivid mental images as an aid for memorizing lists of words. They were shown examples using both concrete and abstract words and practiced using a mental imagery strategy on a series of short recall tasks. The imagery training was intended to improve participants' memory strategies by teaching them to associate the to-be-remembered items with concrete, vivid, and meaningful or bizarre mental images, all qualities that should improve subsequent memory performance (Baddeley &

Andrade, 2000; Nelson & Schreiber, 1992; Paivio, 1965; Paivio, Walsh, & Bons, 1994; West & Holcomb, 2000).

We hypothesized that the participants in the control group would not exhibit any significant differences in performance between the pre- and post-training baseline tasks. For the participants in the imagery training group, we expected that using a mental imagery memory strategy would lead to a general improvement on all conditions within the recognition memory task after training. In addition, since strategy use has been shown to improve performance on WM tasks (McNamara & Scott, 2001; St. Clair-Thompson et al., 2010; Turley-Ames & Whitfield, 2003), we hypothesized that the imagery training group's performance on the WM baseline tasks would also improve after training. For the WM training group, we expected to see near transfer in which participants improved their performance on the WM baseline tasks after training. We also predicted that there would be far transfer of the WM training to the recognition memory task. Specifically, if WM training improves WM capacity, we would expect to see higher performance for the once-presented and repeated words, which may be encoded better after spending more time in WM (Braun & Rubin, 1998). We also predicted that an improved WM span might reduce the size of the spacing effect by making the processing of the long lag repetitions more similar to the processing of the short lag repetitions.

Method

Each participant in the experiment completed tasks over the course of a five-week period. During the first week, participants completed pre-training baseline memory tasks that included a verbal WM task (listening span), a spatial WM task (rotation span), and a verbal recognition memory task. During the next three weeks, participants completed memory training sessions that differed based on the training group to which they were assigned. Participants assigned to the mental imagery training group completed three training sessions (one per week) and participants in the WM training group completed 14 training sessions (4-5 per week) during the three week training period. Participants assigned to the control group did not complete any tasks during the training period. At the end of the training period, all participants completed the same baseline tasks for a second time. Each of the baseline and training tasks is described in detail below.

Participants

Eighty-six participants recruited from the employee population of Sandia National Laboratories participated in this experiment and were paid for their time. All were right-handed, had no early exposure to languages other than English, and had no history of neurological disease or defect. Participants were assigned semi-randomly to one of the three training groups (efforts were made to balance the three groups based on age and gender). Eight participants dropped out of the study before completing all of the sessions and four additional participants failed to follow instructions and were excluded from the data analysis. Of the remaining 74 participants, 25 (12 female) were in the control group, 24 (10 female) were in the imagery training group, and 25 (13 female) were in the WM training group. The mean age for all of the participants was 37 (range 18-63). The mean ages for each group were 37 for the control group (range 18-61), 39 for the imagery training group (range 18-63) and 35 for the WM training group (range 20-63). As the demographics of the Sandia employee population are quite different from college student populations, the distributions of age and educational background for the participants in this study are shown in Figure 1.

Baseline Tasks

Listening Span Task. Based on Daneman & Blennerhassett (1984), the listening span task required participants to recall a sequence of symbols in the order in which they were presented. The presentation of the symbols was interleaved with the auditory presentation of sentences. Participants had to indicate whether the sentences made sense or not. Participants practiced the two tasks separately, and then performed both in the dual task phase.

***Design*.** The dependent variable in the listening span task was the total number of symbols recalled in the correct order during a dual-task test phase. Participants had to maintain a high level of performance in the sentence judgment task, (at least 85% accuracy) in order for their data to be included in the analysis.

***Materials*.** Materials for the memory task were nine black Wingdings symbols in size 24 font presented against a white background:

🔔 ☎ 🕯 ☺ ✂ 💣 ⚐ ✎ ☠

The secondary task was comprised of 110 spoken sentences, half of which were sensible ("They gave the waiter a tip even though he was rude"), and half of which were not ("The children were summer and wanted their parents to come home"). The same speaker was used for all sentence recordings.

Procedure. The listening span task had three phases, the first of which was a memory task involving the sequences of symbols. Participants saw a series of symbols that were presented for 1,000 ms each in the center of a computer screen. They were then shown a recall screen that displayed all of the symbols and were asked to select the symbols in the order in which they had appeared. Participants could edit their selections and insert "blanks" into the sequence in place of symbols that they could not recall. When the participants were satisfied with the sequence of symbols in their responses, they clicked "next." Following each response screen, participants saw feedback on their performance (i.e., "you recalled X of Y items correctly").

The second task required participants to judge the sensibility of sentences. Sentences were presented using headphones, which participants adjusted to a comfortable volume. During sentence presentation, participants received instructions to click the mouse once they could tell whether the sentence was made sense or not, at which point the sentence stopped playing and they responded by clicking either the yes or no button on the screen. Following this response, feedback on accuracy appeared on the screen.

During the dual task phase, participants saw a new symbol after judging each sentence. The recall screen appeared after a sequence of 4 to 8 symbols had accrued. Each participant saw two sequences of 4 and 5 symbols, and three sequences of 6, 7, and 8 symbols. The different sequence lengths were randomly ordered for each participant.

Rotation Span Task. Based on Shah & Miyake (1996), this task required participants to recall sequences of arrows of varying length and orientation. The presentation of the arrows was interleaved with the presentation of letter characters. Participants had to make a judgment as to whether the letters appeared normally or backwards. After each block, participants were asked to recall the sequence of arrows. As in the listening span task, participants practiced the two tasks separately, and then performed both in the dual task phase.

Design. The dependent variable in the rotation span was the total number of arrows recalled during a dual-task test phase. Participants had to maintain a high level of performance in the rotation judgment task (at least 85% accuracy) in order for their data to be included in the analysis.

Materials. For the memory task, the item set was comprised of pictures of long and short arrows at eight orientations, (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°). The secondary task used five letters, (R, L, J, G, and F), at eight different orientations, (45°, 90°, 135°, 180°, 225°, 270°, 315°), with both normal and backwards versions (flipped along the vertical axis). All of the stimuli where white and were presented on a black background.

Procedure. On each trial, participants saw a sequence of rotated letters, each of which was followed by an arrow. For each letter, the participants had to press a key on the keyboard to indicate whether the letter was presented normally or backwards. The letter remained on the screen until the participant made a response. Then an arrow was presented for 1000 ms. The trials varied in length and contained between two and five arrows. After the last arrow was presented, participants were asked to recall the sequence of arrows that they had seen in that trial. The recall screen showed all 16 possible arrows (long or short arrows at each of eight orientations). Participants clicked on the arrows to indicate which arrows had appeared in the previous sequence, in the order that they appeared. As in the listening span task, participants could edit their response and insert blanks in the place of arrows that they had forgotten. When the participants completed their responses for the current trial, they clicked the “next” button to advance to the next trial.

Recognition Memory Task. In the recognition memory task, participants were shown a list of common English nouns and were asked to memorize them for a subsequent recognition test. Some of the nouns were repeated at short (one intervening item) or long (nine intervening items) lags during the study blocks, while other study items were quizzed at short or long lags during the study blocks. All of the studied items were subsequently tested in a recognition memory test, intermixed with an equal number of new, unstudied items.

Design. The critical variables in the recognition memory task were the conditions in which the words were studied (studied once, studied twice, or studied and then quizzed during the study block) and the lag between the study-study and study-quiz repetitions (one or nine intervening items). Both of these variables were manipulated within-subjects. The dependent variable for the behavioral test was yes/no recognition for the studied words in each condition and at each lag.

Materials. The recognition memory task used a list of 1344 words, all of which were common English nouns. The average length of the nouns was 5 letters and their average frequency was 55.67 (based on the Kucera and Francis [1967] norms included in Balota et al., 2002). The words were assigned to counterbalanced experimental lists such that every word appeared in every study and test condition across lists.

The experimental lists were divided into six study-test blocks with equal numbers of each item type in each block. The words were placed in a pseudorandom order within the blocks such that no more than three items in the same condition appeared in sequence. Within each study-test block, there were 28 words that were studied once, 14 words that were studied twice with a short lag between repetitions, 14 words that were studied twice with a long lag between repetitions, 14 words that were studied and then quizzed after a short lag, and 14 words that were studied and then quizzed after a long lag. In addition to the studied items, there were 28 words that served as new items for the quizzes within the study blocks (these words were quizzed but had not been studied) and 84 words that served as new, unstudied items in the subsequent recognition test. In total, each study block contained 112 study words (including repeated study words) and 56 quizzed words. Each test block contained 168 test words, half of which had been studied and half of which were new.

Three of the study-test blocks were presented to each participant during the pre-training baseline session and the other three were presented during the post-training session. The placement of the blocks (pre- or post-training) was counterbalanced across participants.

Procedure. The participants were instructed that they would be tested on their memory for a list of study words. Throughout each task, a fixation cross was shown in the center of the screen. Prior to the presentation of each word, a yellow or red dot appeared on the screen immediately above the fixation cross. Participants were instructed that the yellow dot indicated that the next word was a study word and that they should silently read that word and try to remember it for later. They were told that the red dot indicated that the next word was a quiz word, and that following the word they should press a button to indicate whether or not they had studied that word earlier in the session. The study or quiz word was presented 600-800 ms after the dot disappeared and remained on the screen for one second. The words were presented immediately above the fixation cross in white 48-point Arial font on a black background. If the word was a quiz word, it was followed by a red question mark that remained on the screen until the participant pressed a response button. Participants pressed one of two buttons on a game controller, labeled “yes” and “no” to indicate whether or not they remembered studying that word.

At the end of each study block, participants took a short break before beginning the test block. All of the words in the test block were presented in the same way as the quizzed items from the study block. Each test word was preceded by a red dot and followed by a question mark. While the question mark was on the screen, participants pressed the “yes” or “no” button to indicate whether or not they remembered studying that word during the study block. It took participants approximately 10 minutes to complete each study block and 12 minutes to complete the corresponding test block.

Training Tasks

Mental Imagery Training. In the three weeks in between the pre-training and post-training baseline sessions, 24 of the participants completed three memory training sessions in which they practiced using a mental imagery strategy to remember word lists for a free recall test. The training sessions became more difficult as the participants progressed by using longer word lists, shorter encoding times, and more words with low imagability.

Materials. The memory tests used in the mental imagery training consisted of 168 nouns. Care was taken to ensure that none of the words used in the training sessions appeared in any of the pre-training or post-training baseline tasks. Of the 168 nouns, 49 had low imagability (ratings below 400 in the norms included in the MRC Psycholinguistic Database, Wilson, 1988) and the remainder had high imagability (ratings above 550).

Procedure. For all of the mental imagery training sessions, participants had the option of completing the training in the lab or by logging in to a lab computer via a remote desktop connection from their own computer. The training consisted of three sessions and participants were asked to complete the sessions once a week for three weeks. Each session took approximately half an hour to complete. During the training, the participants were then given examples of mental imagery. The examples, which included both concrete and abstract concepts, explained that creating detailed and unusual mental images could be helpful for remembering information. After a short practice session in which participants were asked to generate and describe mental images for a short list of words, the training provided examples of grouping several mental images into one scene to increase their memorability. The participants were then asked to practice the mental imagery strategy by memorizing two lists of words, each of which was followed by a recall test. During the first practice list, the participants controlled the presentation of the study words and could view each word for as long as they wanted. For the second practice list, each word was presented for three seconds. Each practice list contained 10 words and participants had 10 chances to enter the words during the recall test. They received feedback after entering each word to indicate whether or not each entry was correct. After the memory test, the participants were asked to describe the mental images that they had generated for the word list and to rate the effectiveness of the mental imagery strategy.

In the second and third training sessions, participants saw a brief review of the examples of mental imagery that were presented in the first session and were then asked to practice the mental imagery strategy while completing memory tests with the same format as the tests used in the first training session. As the training progressed, the study lists became longer and included more words with low imagability. The encoding time per word also decreased to two seconds. The structure of each study list is shown in Table 1.

Working Memory Training. The 25 participants in the WM training group were trained on two tasks, the adaptive N-back task and the symmetry span task. Participants were loaned a laptop containing the two tasks and were asked to do each task once on every business day for three weeks. They were told that they could skip the training on one day of their choice, for a total of 14 sessions.

Adaptive N-back Task. In the adaptive n-back task, single letters appeared sequentially on the screen and participants were required to indicate with a button press whether the current letter had appeared N items previously or not. For example, if subjects were shown the sequence A-B-C-B in a 2-back task, they would indicate that the second “B” was a target because it matched the letter that had appeared two letters back. They would respond ‘non-target’ to the other items in the sequence. The version of the task used in the current experiment (see also Novick et al., submitted) included lure trials. The lures were letters that appeared before (N-1) or after (N+1) the nth-back item (Kane, Conway, Miura, & Colflesh, 2007; see also Gray et al., 2003). For example, in the sequence A-B-A-C-D in a 3-back task, the second A is a lure (an N-1 lure) because it repeats a letter that appeared previously but not the correct number of letters back for the 3-back task. Subjects would have to resolve the interference of these familiar but non-target stimuli. Participants were presented with three lure conditions that corresponded with three levels of difficulty (no lures, N+1 lures only and both N+l and N-1 lures). The difficulty level of the task changed based on the participant’s performance. When participants achieved at least 85% accuracy at the current level (of n and lure difficulty level), difficulty was increased, first by increasing lure difficulty and then by increasing the n. If accuracy fell below 65%, task difficulty decreased, first by decreasing lure difficulty and then by decreasing n. Task difficulty, therefore, represented both the value of n and the lure difficulty level.

Symmetry Span Task. The symmetry span task required participants to remember the locations of a sequence of blocks that appeared in a 4x4 grid, in the order in which they were presented. The presentation of the blocks was interleaved with the presentation of a design on a different grid. Participants had to determine if the design was symmetrical across the vertical axis. Participants practiced the two tasks separately and then performed both in the dual-task phase. At the end of a series of these presentations, participants reported the remembered blocks by clicking on their positions, in order, on a blank grid.

The difficulty of this task was adjusted by changing the number of blocks that the participant needed to remember. The length of each series of memory items was determined by the participant's performance on the last set of judgments. Performance was re-evaluated after every four memory responses. If the participant got three or more correct, the sequence length increased by one. Conversely, if performance fell below two correct, the sequence length decreased by one. All participants started with a sequence length of three blocks.

Memory Strategy Survey. At the end of their participation in the study, the participants were asked to complete a follow-up questionnaire about their use of various memory strategies. Fifty-six of the participants completed the survey (18 from the control group, 19 from the mental imagery training group, and 19 from the WM training group). The first part of the survey asked participants to describe their memory strategy for each task. The second part gave examples of different memory strategies (mental imagery, generating sentences or stories, linking items to one another, rehearsal and self-quizzing) and participants were asked if they had used those strategies on the pre- and post-training memory tasks. Participants also completed two questionnaires about their use of mental imagery, the Object-Spatial Imagery Questionnaire (OSIQ; Blajenkova, Kozhevnikov & Motes, 2006) and the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973).

Results

Mental Imagery Training

Twenty-three of the 24 participants completed all three of the imagery training sessions and one participant completed only two of the training sessions. The average number of words recalled by the participants remained fairly consistent across the 14 memory tests used in the imagery training session, even as the encoding task became more difficult (longer word lists, shorter encoding times, more abstract words). Participants recalled an average of 9.04 words on the first memory test and an average of 8.14 words on the last memory test. As the memory tests became more difficult, participants reported that it became more difficult to create mental images for the word lists and they felt that the imagery strategy was less effective for the more difficult lists. The participants' ratings of how easy it was to create mental images (where 1 was easy and 5 was difficult) increased from an average of 2.96 for the first memory test to an average of 4.0 for the last, most difficult test. Similarly, the participants' ratings of the effectiveness of the memory strategy (where 1 is not effective and 5 is very effective) decreased from an average of 4.17 for the first memory test to an average of 2.5 for the last memory test.

Working Memory Training

Twenty-four of the 25 participants in the WM training group completed at least 12 of the 14 WM training sessions, and one participant completed 9 of the training sessions. The participants who completed at least 12 of the WM training sessions were included in the analysis of the WM training. On average, the participants' performance improved across the training sessions for both training tasks. During the first training session, the participants had an average N-back level of 1.81 and an average symmetry span difficulty level of 3.77. On the 12th training session, the participants' average N-level was 4.23 and their average symmetry span difficulty level was 5.43. However, there was a great deal of variability across participants. The average N-level achieved by each participant on the 12th training session ranged from 1 to 8.83. Similarly, the average level of difficulty achieved by each participant for the 12th session of the symmetry span task ranged from 3.18 to 7.33. Figure 2 shows the changes in performance across the WM training sessions.

Baseline Memory Tasks

While participants generally improved their performance on the tasks on which they were trained, the key question was whether or not their training would affect their performance on untrained memory tasks. To address this question, we compared the three training groups' changes in performance on the three pre- and post-training baseline tasks.

Rotation Span Task. Two participants were excluded from the analysis of the rotation span task (one from the control group and one from the WM training group) due to failure to complete the post-training test. The mean accuracy for the remaining participants in each training group are shown in Table 2. Paired t-tests were used to assess each group's change in performance between the pre-training and post-training sessions. All three training groups performed significantly better during the post-training session (all t's > 3.62, p's < 0.01). However, a one-way ANOVA comparing the change in performance across all three training groups showed that there were no significant differences between the groups ($F(2, 69) = 0.49$, $p = 0.61$).

Listening Span Task. Five participants (four from the control group and one from the WM training group) were excluded from the analysis of the listening span task due to a problem with the presentation of the sound files during the pre-training session. The average total scores from the remaining participants in each training group are shown in Table 3. Paired t-tests were used to asses each group's change in performance between the pre-training and post-training sessions. All three groups performed significantly better during the post-training session (all t's >1.87, p's < 0.04). However, a one-way ANOVA showed that the three groups were not significantly different in terms of how much their performance improved ($F(2, 66) = 1.69$, $p = 0.19$).

Recognition Memory Task. The participants' recognition memory performance (average proportion correct for each condition) is shown in Table 4. For analysis of the pre-training session, the performance was collapsed across training groups. The average performance of all participants on the pre-training recognition test is shown in Figure 3. Paired t-tests showed that, as predicted, the participants' memory performance was significantly better for words that were repeated or quizzed relative to words that were studied only once (all t's > 10.31, all p's < 0.001). In addition, there was a spacing effect for both the repeated and the quizzed study items such that participants had significantly better memory performance for long lag items than for short lag items ($t(73) = 2.72$, $p < 0.01$ for repeated words; $t(73) = 5.13$, $p < 0.001$ for quizzed words).

The crucial comparison for examining the effects of the memory training techniques was the difference between pre-training and post-training recognition memory performance for the three training groups. The differences for each group (post-training accuracy minus pre-training accuracy on each of the memory test conditions) are shown in Figure 4. Paired t-tests were used to assess each group's change in performance for each test condition. The t-tests showed that the control group's performance in the pre-training and post-training sessions did not differ significantly for any condition (all t's < 1.12, all p's > 0.27). The participants in the mental imagery training group performed significantly better in the post-training session for the once-presented words ($t(23) = 1.80$, $p = 0.04$), the short-lag repeated words ($t(23) = 2.40$, $p = 0.01$) and the short-lag quizzed words ($t(23) = 2.17$, $p = 0.02$). The participants in the WM training group performed significantly *worse* in the post-training session relative to the pre-training session for the long-lag repeated items ($t(24) = 2.81$, $p < 0.01$), the short-lag quizzed words ($t(24) = 2.26$, $p = 0.02$) and the long-lag quizzed words ($t(24) = 4.55$, $p < 0.01$). Their performance was marginally worse for the once-presented words ($t(24) = 1.65$, $p = 0.06$).

For each test condition, a one-way ANOVA was used to compare the change in performance (post-training accuracy minus pre-training accuracy) across the three groups.

Although the three training groups' change in performance did not differ significantly for correct rejections of new, unstudied items ($F(2, 71) = 0.30$), there were significant differences in the participants' change in hit rates in every other condition (once-presented words, $F(2, 71) = 3.50$, $p = 0.04$; short-lag repeated words, $F(2, 71) = 3.33$, $p = 0.04$; long-lag repeated words, $F(2, 71) = 3.60$, $p = 0.03$; short-lag quizzed words, $F(2, 71) = 4.23$, $p = 0.02$; long-lag quizzed words, $F(2, 71) = 8.12$, $p < 0.01$).

Memory Strategy Survey

The participants' responses to the initial, free-response portion of the memory strategy survey were used to categorize their self-reported strategy into one of seven categories: semantic, imagery, semantic + imagery, rehearsal, semantic + rehearsal, other or none. Semantic strategies included techniques such as forming associations between items, inventing sentences or stories, or trying to relate the items to personal experiences. The two participants categorized as "other" both reported that they tried to remember the first letter or first syllable of the words. Representative examples of each category are shown in Table 5.

A breakdown of the number of participants in each group who reported using each strategy is shown in Figure 5. As expected, the majority of the participants in the imagery training group reported using mental imagery. Notably, no one in the imagery training group reported that they used a strategy that involved rehearsal although rehearsal-based strategies were reported by participants from both of the other groups.

To examine the impact of memory strategy on recognition memory performance, we assessed the participants' average hit rate across all conditions of the recognition memory task. The "Other" category was combined with the "None" category due to the small number of participants in that category and the nature of their reported strategies. The results are shown in

Figure 6. A one-way ANOVA showed that memory performance differed significantly across the different memory strategies ($F(5, 50) = 10.22$, $p < 0.01$). Participants who reported using an imagery strategy performed best, and t-tests showed that their performance was significantly higher than that of the participants who reported using a semantic strategy, rehearsal strategies (alone or in combination with a semantic strategy), or no strategy (all t's > 2.57, all p's < 0.02). Participants who used imagery in combination with a semantic strategy performed better than participants who used repetition strategies or no strategy (all t's > 3.48, all p's < 0.01), but their performance was not significantly different from that of participants who reported using semantic or imagery strategies alone ($t(24) = 1.80$, $p = 0.08$ and $t(17) = 0.28$, respectively). The participants who reported using a rehearsal strategy, rehearsal combined with a semantic strategy, or no strategy had the lowest memory performance. Those three groups did not differ significantly from one another (all t's < 1.47).

Discussion

The results of this study add to the growing body of literature that suggests that WM training does not necessarily transfer to other memory tasks. Although the participants in the WM training group did improve their performance on the trained tasks, we failed to find near transfer to other WM tasks. The participants in the WM training group did perform better on the untrained baseline WM tasks after training, but their performance did not improve any more than that of the participants in the control groups, indicating that the improvement was not related to the WM training. This study also revealed that WM training can have negative effects on other types of memory tasks. The results showed that the participants in the WM training group performed worse on the recognition memory task after training, while participants in the no-contact control group maintained the same level of performance and participants in the mental imagery training group improved their performance.

One of the key differences between the mental imagery training group and the WM training group is that former was trained to use a memory strategy and the latter was not. Differences in use of memory strategies between the two groups could account for their performance changes on the recognition memory test. Presumably, each participant approached the baseline recognition memory task with some sort of memory strategy. The participants in the no-training control group would be likely to use the same strategy for the task both times they completed it, leading to similar performance in both instances. The participants who practiced using a mental imagery strategy tried to apply that strategy to recognition memory task and their performance improved. The WM training group's decline in performance suggests that, for a task in which a good memory strategy is helpful, WM training may lead participants to adopt less effective memory strategies.

Prior research on the effects of strategy use on WM task performance has indicated that asking participants to use a rehearsal strategy can improve WM span, particularly for participants with low span scores prior to strategy training (Turley-Ames & Whitfield, 2003). Other strategies, such as imagery and semantic elaboration were not as helpful, perhaps because enacting those strategies placed too much of a burden on limited WM resources. McNamara and Scott (2001) also found that participants who completed WM and short-term memory tasks tended to use a rehearsal strategy unless they were trained to use a semantic chaining strategy. In the present study, the intensive WM training may have led participants to favor rehearsal strategies in order to improve their performance during the training sessions. However, rehearsal strategies were not very effective for the recognition memory task, which had over 100 words per block. The results of the memory strategy survey show that use of a rehearsal-based strategy (or no strategy at all) was associated with lower performance on the recognition memory test. Participants in the WM training group were more likely than participants in the other groups to report using one of these less effective strategies. The survey responses support the idea that the WM training may have led participants to adopt strategies that were ill-suited to the recognition memory task.

Development of an effective (or ineffective) memory strategy could also explain the difference between our results and those of Rudebeck and colleagues (2012), which found improvements in recognition memory performance after WM training. Their study used a spatial WM training task in which participants were presented with greyscale images of real-world scenes in one of 8 locations. The recognition memory task used images that were similar to the ones used in the training task. If the participants developed strategies for remembering greyscale images of scenes during training, that same strategy would benefit them during the recognition memory test. In the present study, types of items used during training were different from those used in the recognition memory test, so participants could not benefit from any item-specific memory strategies that they may have practiced during the training. More general strategies such as rehearsal that may have been useful during training were not well-suited to the demands of the verbal recognition memory test.

Memory Strategies and WM Measures

As past studies have noted, measures of WM capacity are seldom free from strategy use. McNamara and Scott (2001) reported that in the absence of instructions regarding strategy use, 55 out of 60 participants utilized a strategy on a WM baseline task, providing evidence that a strategic approach to performing WM tasks is the norm rather than the exception. Similarly, Turley-Ames and Whitfield (2003) instructed to participants to use rehearsal, visual imagery, or a semantic strategy during the performance of an operation span task, and found an increase of WM scores as a result of each strategy. Since strategy use influences task performance as well as the predictive validity of WM span scores for cognitive abilities, an accurate assessment of WM capacity necessitates control over variability in strategy use (Turley-Ames & Whitfield, 2003).

While the goal of assessing WM capacity is to predict higher-order cognitive functioning, the goal of WM training is to enhance higher-order cognitive functioning. WM training is intended to transfer to tasks beyond the trained tasks so that it can provide more general cognitive benefits to the trainees. In order to foster transfer to other tasks, researchers have used adaptive training tasks to minimize the formation of task-specific memory strategies during training (e.g. Jaeggi et al., 2008). If task-specific strategies are minimized, improvements during training should reflect enhancement of the underlying construct rather than task-specific improvement. However, the present study indicates that adaptive WM training tasks may lead people to abandon memory strategies that would be beneficial when applied to other types of memory tasks. When focusing on WM enhancement, it may be more useful to think of memory strategies as complementary interventions. Development of task-specific strategies remains a critical concern, but rather than discouraging the formation of strategy use via adaptive tasks, it may be beneficial to explicitly teach participants strategies that are themselves applicable to a wide variety of tasks.

The current study presents a cautionary tale about unintended consequences arising from cognitive training. Even if participants improve on the trained task, the training may impact their performance on untrained tasks in unforeseen ways. If intensive cognitive training changes trainees' use of strategies, it could impact on performance on many other tasks. This is particularly concerning if the mental strategies that are reinforced during training are maladaptive for other common tasks. Many studies have investigated near and far transfer of WM training to WM and fluid intelligence tasks, but few studies have examined the impact of WM training on other types of memory performance. In this study, we found an example of negative transfer in which WM training decreased performance on a recognition memory task. Given the broad popularity and time-consuming nature of WM training, its effects on other types of memory that are crucial for learning, and the interplay between WM training and memory strategies are clearly areas that warrant additional research.

References

Aamodt S., & Wang, A. (2007, November 8). Exercise on the brain. *The New York Times*. Retrieved from http://www.nytimes.com/2007/11/08/opinion/08aamodt.html?_r=1

Allan, K., Rugg, & M. D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia, 35*, 387-397.

Alloway, T. P., Bibile, V., & Lau, G. (2013). Computerized working memory training: Can it lead to gains in cognitive skills in students? *Computers in Human Behavior, 29*, 632-638.

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context-availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers, 31,* 578-602.

Anderson, J. E., & Holcomb, P. J. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and Language, 94*, 200-216.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4,* 829–839.

Baddeley, A. D., & Andrade, J. (2000).Working memory and the vividness of imagery. *Journal of Experimental Psychology, General, 129*, 126-145.

Bahrick, H.P., & Hall, L.K. (2005). The importance of retrieval failure to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language, 52*, 566-577.

Balota, D.A., Cortese, M.J., Hutchison, K.A., Neely, J.H., Nelson, D., Simpson, G.B., & Treiman, R. (2002). The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. http://elexicon.wustl.edu/, Washington University.

Becker, J. T. & Morris, R. G. (1999).Working Memory(s). *Brain and Cognition, 41,* 1-8.

Benjamin, A. S. (2008). Memory is more than just remembering: Strategic control of encoding, accessing memory, and making decisions. In Ross, B.H., Benjamin A.S. (eds.) *The Psychology of Learning and Motivation: Skill and Strategy in Memory Use. Vol 48*, pp. 175-223. Academic Press, London.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127*, 55-68.

Bentin, S., & Moscovitch, M. (1990). Psychophysiological indices of implicit memory performance. *Bulletin of the Psychonomic Society, 28(4),* 346-352.

Bentin, S., & Peled, B. S. (1990). The contribution of task-related factors to ERP repetition effects at short and long lags. *Memory and Cognition, 18*, 359-366.

Besson, M., & Kutas, M. (1993). The many facets of repetition: Cued-recall and event-related potential analysis of repeating words in same versus different sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1115-1133.

Besson, M., Kutas, M., & Van Petten, C. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience, 4*, 132-149.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimatnura (Eds.), *Metacognition: Knowing about knowing*, pp. 185-205. Cambridge, MA: MIT Press.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and Illusions. In Gopher D., Koriat A. (eds.) *Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application*. pp. 435-459. MIT Press, Cambridge, MA

Blajenkova, O., Kozhevnikov, M., & Motes, M. A. (2006). Object-Spatial imagery: A new self-report imagery questionnaire. *Applied Cognitive Psychology, 20*, 239-263.

Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval, and time in working memory: Evidence from once-presented words. *Memory, 6,* 37-65.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience, 2,* 51-61.

Carpenter, P. A., Just, M. A., & Shell, P. (1990).What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*, 404–431.

Carretti, B., Borella, E., Fostinelli, S., & Zavagnin, M. (2013). Benefits of training working memory in amnestic mild cognitive imapairment: Specific and transfer effects. *International Psychogeriatrics, 25*, 617-626.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory.In Anderson, J. R. (Ed.), *Cognitive skills and their acquisition* (pp. 141-189). Hillsdale, NJ: Erlbaum.

Chooi, W., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence, 40*, 531-542.

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 & Review, 17,* 193–199.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. Psychonomic Bulletin & Review, 12, 769-786.

Dahlin, E., Neely, A. S., Larsson, A., Backman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320,* 1510-1512.

Daneman, M., & Blennerhassett, A. (1984). How to assess the listening comprehension skills of prereaders. *Journal of Educational Psychology, 76*, 1372 -1381

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading.*Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods, 134*, 9-21.

Engle, R. W., Cantor, J. & Carullo, J. J. (1992). Individual difference in working memory for comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning Memory, and Cognition, 18,* 972-992.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

Ericsson, K., & Delaney, P. (1998).Working memory and expert performance. In Logie, R., & Gilhooly, K. (Eds.), *Working memory and thinking: current issues in thinking and reasoning*(pp. 93-114). Hillsdale, NJ: Erlbaum.

Federmeier, K. D., & Benjamin, A. S. (2005). Hemispheric asymmetries in the time course of recognition memory. *Psychonomic Bulletin & Review, 12 (6),* 993-998.

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.

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology, 24*, 827-836.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., &Perrig, W.J. (2008). Improving fluid intelligence with training on working memory. *PNAS, 105*, 6829-6833.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66-71.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning.*Science, 319,* 966-968.

Koriat, A., & Bjork, R. A. (2005).Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 187-194.

Koriat, A., Bjork, R.A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133*, 643-656.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). The case for mental imagery. New York City, NY: Oxford University Press.

Kucera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.

Kutas, M., & Federmeier, K. D. (2000). Electophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences, 4, 463-470.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In Gruneberg, M. M., Morris, P. E., Sykes, R. N. (Eds.),*Practical Aspects of Memory* (pp. 625-632). Academic Press, London.

Laszlo, S., & Federmeier, K. D. (2007). Better the DVL you know: Acronyms reveal the contribution of familiarity to single word reading. *Psychological Science, 18*, 122-126.

Luo, Y., Wang, J., Wu, H. R., Zhu, D. M., & Zhang, Y. (2013). Working-memory training improves developmental dyslexia in Chinese children. *Neural regeneration research, 8*, 452-460.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory and Cognition, 29*, 10-17.

Marks, D. F. (1973).Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64*, 17-24.

Marschark, M., & Paivio, A. (1977). Integrative processing of concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior, 16*, 217–231.

Matzen, L. E., & Benjamin, A. S. (2009). Remembering words not presented in sentences: How study context changes patterns of false memories. *Memory & Cognition, 37,* 52-64.

Melby-Lervåg, M. & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270-291.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. Journal of Verbal Learning and Verbal Behavior, 9, 596-606.

Nagy, M. E., & Rugg, M. D. (1987). Modulation of ERPs by word repetition: The effects of inter-item lag. *Psychophysiology, 26*, 431–436.

Nelson, D. L., & Schreiber, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language, 31,* 237–260.

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. (1986). Event-related brain potentials during the initial encoding and subsequent recognition memory of congruous and incongruous words. *Journal of Memory and Language, 25*, 75-92.

Oh, V. M. (1994). The placebo effect: Can we use it better? *British Medical Journal, 309,* 69-70.

Olichney, J. M., Van Patten, C., Paller, K. A., Salmon, D. P., Iragui, V. J., & Kutas, M. (2000). Word repetition in amnesia: Electrophysiological measures of impaired and spared memory. *Brain, 123*, 1948-1963.

Paller, K. A., & Gross, M. (1998). Brain potentials associated with perceptual priming versus explicit remembering during the repetition of visual word-form. *Neuropsychologia, 36,* 559-571.

Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neuropsychological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience, 4*, 375-391.

Paivio, A. (1965). Abstractness, imagery and meaningfulness inpaired associate-learning.*Journal of Verbal Learning and Verbal Behaviour, 4,* 32-38.

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1196–1204.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29*, 633-654.

Pujol, J., Deus, J., Losilla, J. M., Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology, 52*, 1038.

Rohde, T. E., & Thompson, L. A. (2007).Predicting academic achievement with cognitive ability.*Intelligence, 35*, 83-92.

Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 803-814.

Roediger, H.L., Watson, J. M., McDermott, K. B. & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8(3),* 385-407.

Rudebeck, S. R., Bor, D., Ormond, A., O'Reilly, J. X., & Lee, A. C. H. (2012). A potential spatial working memory training task to improve both episodic memory and fluid intelligence. *PLoS ONE 7*(11): e50431. doi:10.1371/journal.pone.0050431

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials.*Psychophysiology, 22*, 642–647.

Rugg, M. D. (1987). Dissociation of semantic priming, word and nonword repetition effects by event-related potentials.*Quarterly Journal of Experimental Psychology, 39A*, 123–148.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition, 18*, 367–379

Rugg, M. D., & Doyle, M. C. (1994). Event-related potentials and stimulus repetition in direct and indirect tests of memory. In Heinze, H. J., Munte, T. F., Mangun, G. R. (eds.) *Cognitive Electrophysiology.* pp. 124-138. Birkhauser, Boston, MA.

Schwartz, B.L., Benjamin, A.S., & Bjork, R.A. (1997). The inferential and experiential bases of metamemory. Current Directions in Psychological Science, 6, 132-137.

Shah P. & Miyake A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4–27.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138,* 628-654.

Singley, M. K., & Anderson, J. R. (1989). *The Transfer of Cognitive Skill*. Cambridge, MA: Harvard University Press.

St. Clair-Thompson, H., Stevens, R., Hunt, A., Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology 30,* 203-219.Smith, M. E., & Halgren, E. (1989). Event-related potentials during lexical decision: Effects of repetition, word frequency, pronounceability, and concreteness. *Electroencephalography and Clinical Neurophysiology, Supplement, 40*, 417–421.

Thompson, T. W., Waskom, M. L., Garel, K. A., Reynolds, G. O., Winter, R., Chang, P., & Gabrieli, J. D. E. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS ONE 8*(5): e63614. doi:10.1371/journal.pone.0063614

Thorndike, E. L., & Woodworth, R. S. (1901).The influence of improvement in on mental function upon the efficiency of other functions (I). *Psychological Review, 8*, 247–261.

Turley, K. J. (1997). *Strategy use and the WM span test*. Paper presented at the 38th annual meeting of the Psychonomic Society, November. Philadelphia, PA.

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49,* 446-468.

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIssac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience, 3*, 131-150.

Van Petten, C., & Senkfor, A. J. (1996). Memory for words and novel visual patterns: Repetition, recognition, and encoding effects in the event-related brain potential. *Psychophysiology, 33*, 491-508.

West, C. W., & Holcomb, P. J. (2001). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*, 1024-1037.

Wilson, M.D. (1988) The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, *20*, 6-11.

MATLAB version 7.11.0.584(R2010b). Natick, Massachusetts: The MathWorks Inc., 2010.

Author Note

Laura E. Matzen, Michael C. Trumbo, Michael J. Haass, Austin Silva and Susan Stevens-Adams, Sandia National Laboratories, and Michael Bunting and Polly O'Rourke, Center for the Advanced Study of Language.

This work was funded by the Sandia National Laboratories Laboratory Directed Research and Developlment Program.

Correspondence concerning this paper should be addressed to Laura E. Matzen, Sandia National Laboratories, PO Box 5800, Mail Stop 1327, Albuquerque, NM, 87185-1327. E-mail: lematze@sandia.gov.

Table 1.

Design of recall tests used in mental imagery training sessions.

Imagery Training Study List	Training Session	Total Number of Study Words	Number of Study Words With Low Imagability	Encoding Time Per Word
1	1	10	1	Self-paced
2	1	10	2	3 seconds
3	2	10	2	3 seconds
4	2	10	3	3 seconds
5	2	10	3	2 seconds
6	2	10	3	2 seconds
7	2	12	3	2 seconds
8	2	12	4	2 seconds
9	3	12	4	3 seconds
10	3	12	4	2 seconds
11	3	14	4	2 seconds
12	3	14	5	2 seconds
13	3	16	5	2 seconds
14	3	16	6	2 seconds

Table 2.

Mean Accuracy on Rotation Span Task for Each Training Group

Training Group	Pre-training Performance	Post-training Performance	Change in Performance (Post – Pre)
Control Group	0.60	0.70	0.10
Mental Imagery	0.57	0.64	0.08
Working Memory	0.57	0.68	0.12

Table 3.

Mean Total Score on Listening Span Task for Each Training Group

Training Group	Pre-training Performance	Post-training Performance	Change in Performance (Post – Pre)
Control Group	34.67	42.76	8.10
Mental Imagery	37.75	43.88	6.13
Working Memory	41.04	44.04	3.00

Table 4.

Mean Accuracy on Recognition Memory Task for Each Training Group

	Condition	Training Group	Pre-training Proportion Correct	Post-training Proportion Correct	Change in Performance (Post – Pre)
Words quizzed during study session	Short lag quizzes	Control Group	0.92	0.94	0.02
		Mental Imagery	0.96	0.98	0.02
		Working Memory	0.92	0.90	-0.01
	Long lag quizzes	Control Group	0.69	0.71	0.02
		Mental Imagery	0.74	0.80	0.06
		Working Memory	0.70	0.63	-0.07
	New words quizzed	Control Group	0.88	0.93	0.05
		Mental Imagery	0.91	0.93	0.02
		Working Memory	0.94	0.93	0.00
Words tested during subsequent memory test	Once-studied words	Control Group	0.84	0.83	-0.01
		Mental Imagery	0.83	0.84	0.01
		Working Memory	0.84	0.82	-0.02
	Short lag repetitions	Control Group	0.47	0.50	0.03
		Mental Imagery	0.57	0.64	0.07
		Working Memory	0.47	0.42	-0.05
	Long lag repetitions	Control Group	0.60	0.62	0.02
		Mental Imagery	0.68	0.76	0.08
		Working Memory	0.61	0.58	-0.03
	Short lag quizzes	Control Group	0.60	0.62	0.02
		Mental Imagery	0.75	0.78	0.03
		Working Memory	0.64	0.55	-0.09
	Long lag quizzes	Control Group	0.65	0.64	-0.01
		Mental Imagery	0.72	0.77	0.05
		Working Memory	0.67	0.61	-0.06
	New words	Control Group	0.68	0.66	-0.02
		Mental Imagery	0.78	0.81	0.03
		Working Memory	0.77	0.65	-0.12

Table 5.

Representative Examples of Each Memory Strategy

Category	Participant's Descriptions
Semantic	"Associated them with something I know or experienced." "I made sentences in my head." "Tried to blend them together in a story."
Imagery	"Tried to create a mental image that represented the word." "Tried to create a vivid image of the word. If it was an object it was much easier than something abstract."
Imagery + Semantic	"String them together in groups that form a little story while thinking of pictures to go with the story." "Tried to form visual images linking the words. Other times made up a story that linked the words." "I tried to associate a mental picture with the word as well as recall the group category e.g. fast, slow."
Rehearsal	"Repetition." "Verbalizing them in my mind." "Saying the words in my head."
Semantic + Rehearsal	"I attempted to repeat them over to myself and make associations with the other words." "I tried to associate the words with things I liked or thought I would remember. I also tried to just memorize words by repeating them as many times as I could."
Other	"Shortened word to letter or single syllable." "I would remember the first letters of the words. Certain words would have meaning and were easily remembered without trying."

None	"I would simply look at the word and think of its meaning." "No real strategy just like absorbing them and add them to a list." "I didn't really use any method. I just paid as much attention to the actual word which gave me a sense of whether I had seen the word or not when tested." "Not much of a strategy… I let the words sort of float to the back of my mind." "Just tried to remember."

Figure Captions

Figure 1. Breakdown of age, gender and education levels of study participants.

Figure 2. Individual and average working memory training results for the N-Back and Symmetry Span tasks.

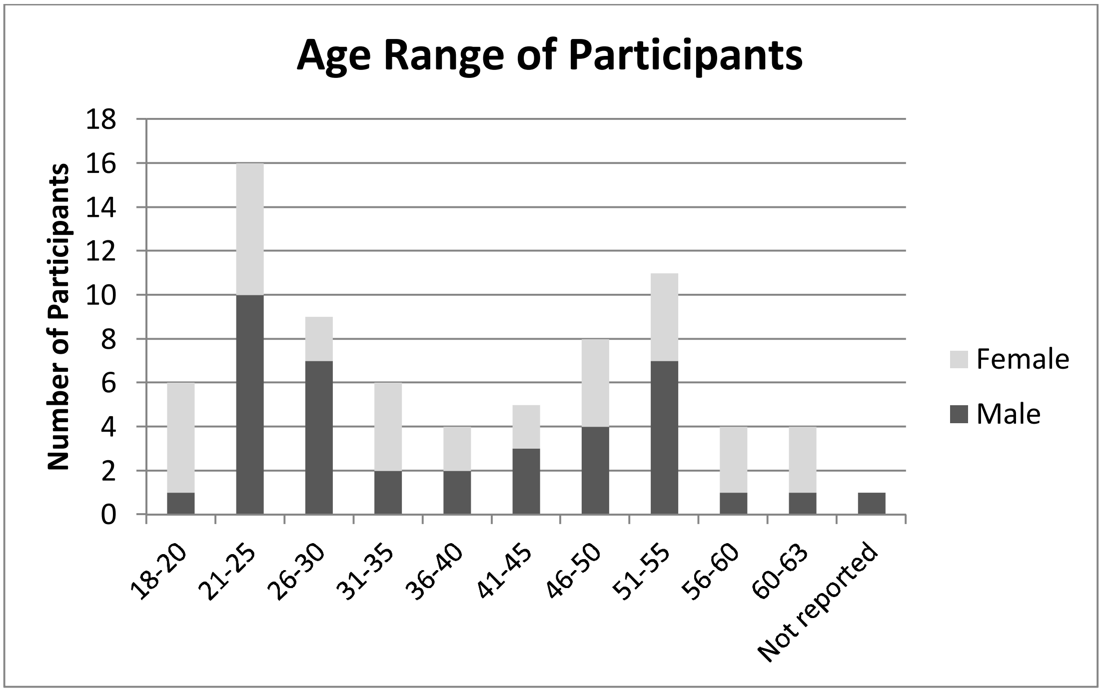
Figure 3. Average proportion of "yes" responses for each condition on the pre-training recognition memory test.

*Figure 4.*Difference between post-training memory performance (% correct) and pre-training memory performance for each condition on the recognition memory test, compared across training groups.

Figure 5. Number of participants in each training group reporting each type of memory strategy.

Figure 6. Average hit rates for participants reporting each type of memory strategy.

Figure 1


Age Range of Participants
Number of Participants
0
2
4
6
8
10
12
14
16
18
18-20
21-25
26-30
31-35
36-40
41-45
46-50
51-55
56-60
60-63
Not reported
Female
Male

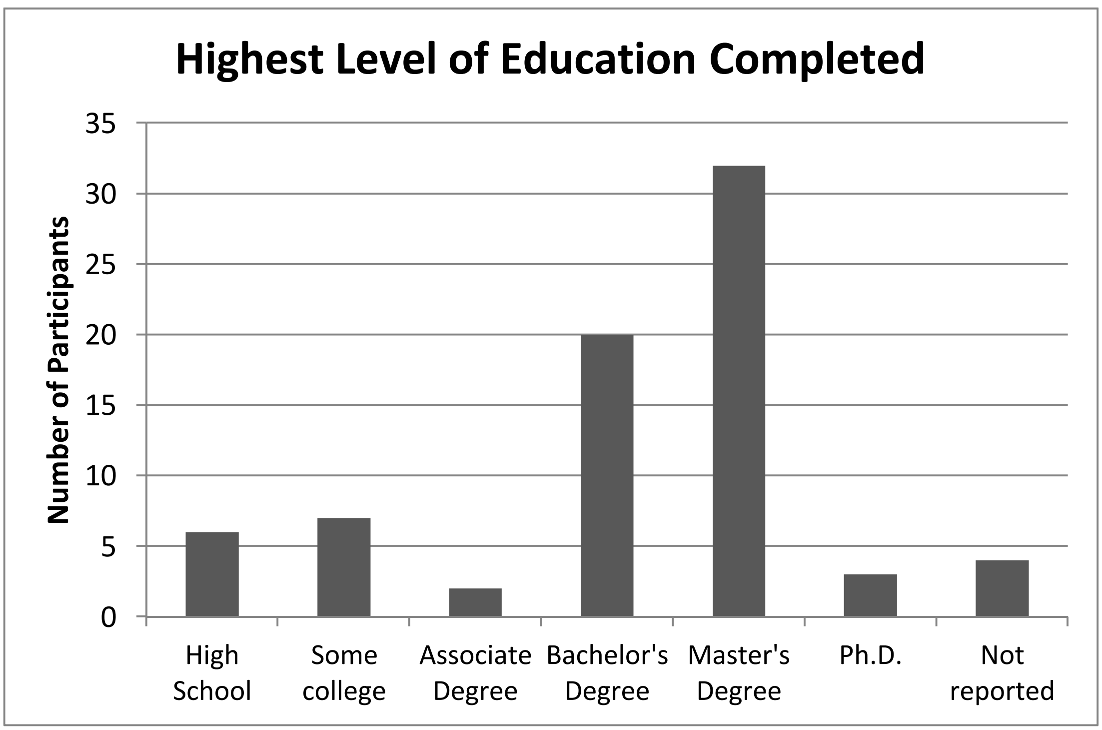

Highest Level of Education Completed
Number of Participants
0
5
10
15
20
25
30
35
High School
Some college
Associate Degree
Bachelor's Degree
Master's Degree
Ph.D.
Not reported

Figure 2.

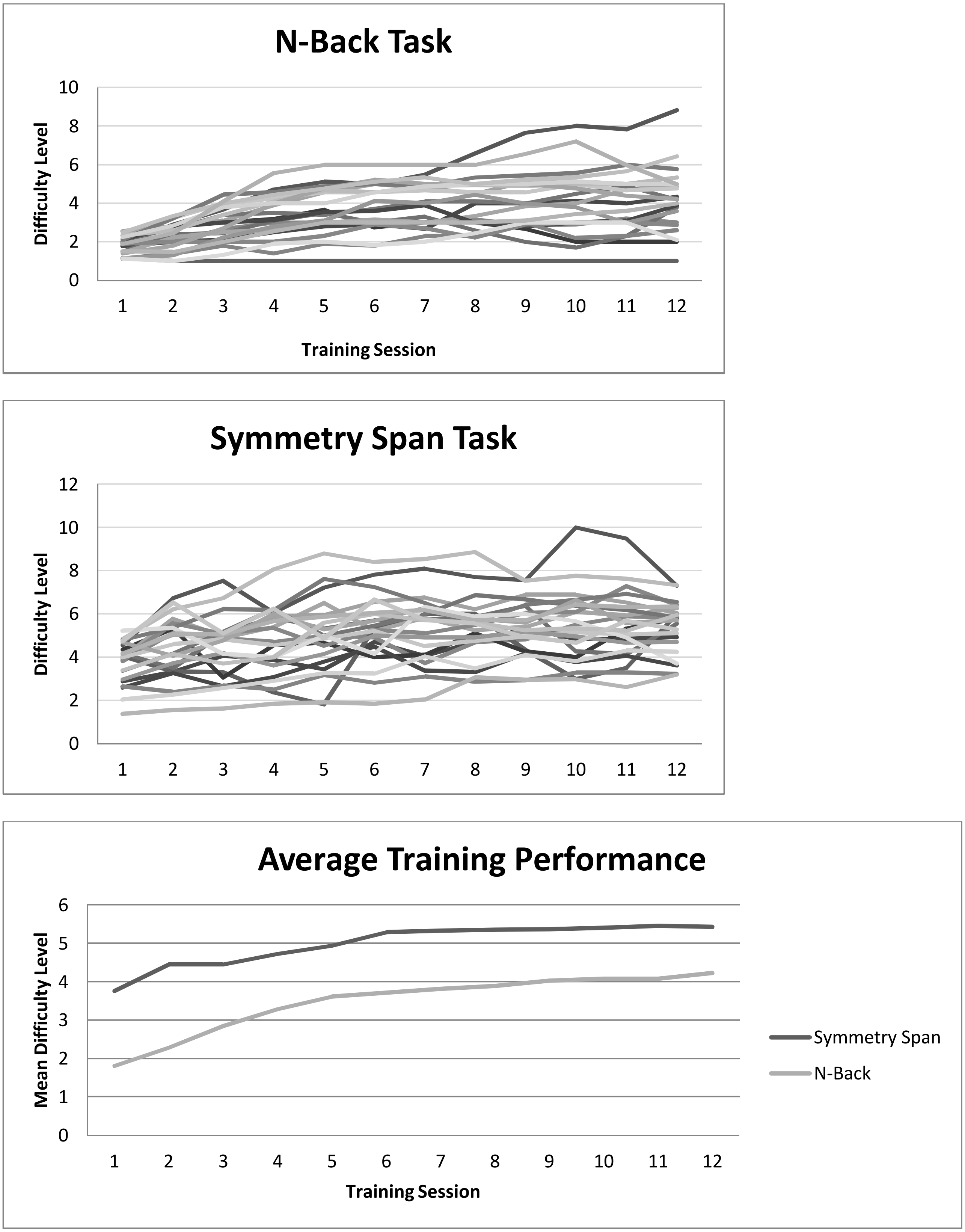

N-Back Task
Difficulty Level
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 10 11 12
Training Session
Symmetry Span Task
Difficulty Level
12
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 10 11 12
Average Training Performance
Mean Difficulty Level
6
5
4
3
2
1
0
1 2 3 4 5 6 7 8 9 10 11 12
Training Session
Symmetry Span
N-Back

Figure 3.

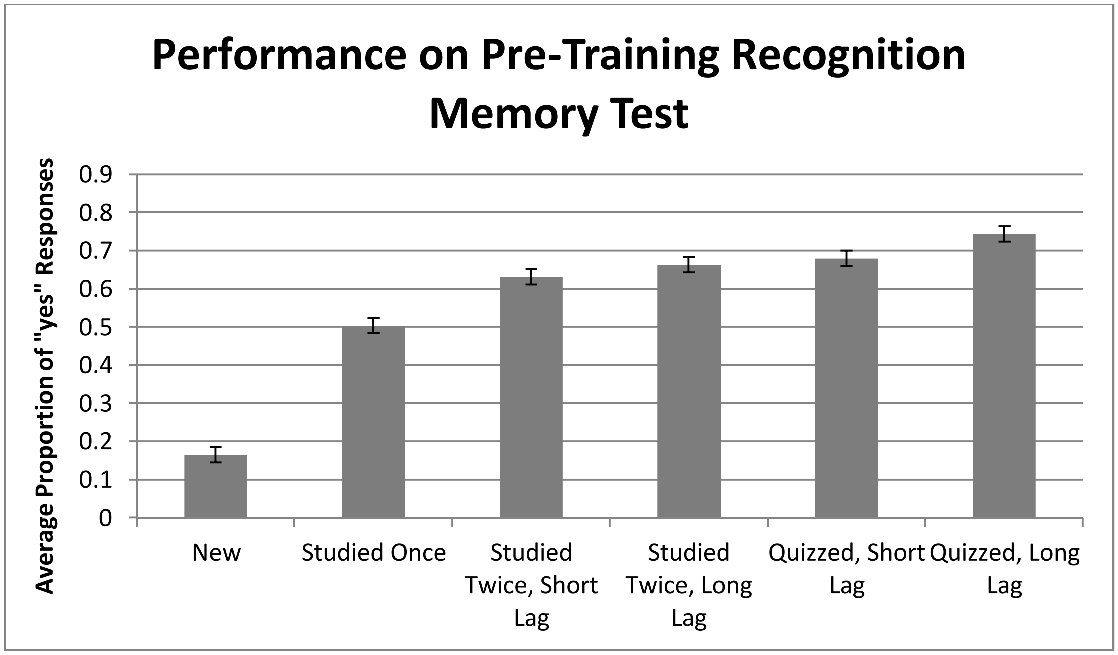

Performance on Pre-Training Recognition Memory Test
Average Proportion of "yes" Responses
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
New
Studied Once
Studied Twice, Short Lag
Studied Twice, Long Lag
Quizzed, Short Lag
Quizzed, Long Lag

Figure 4.

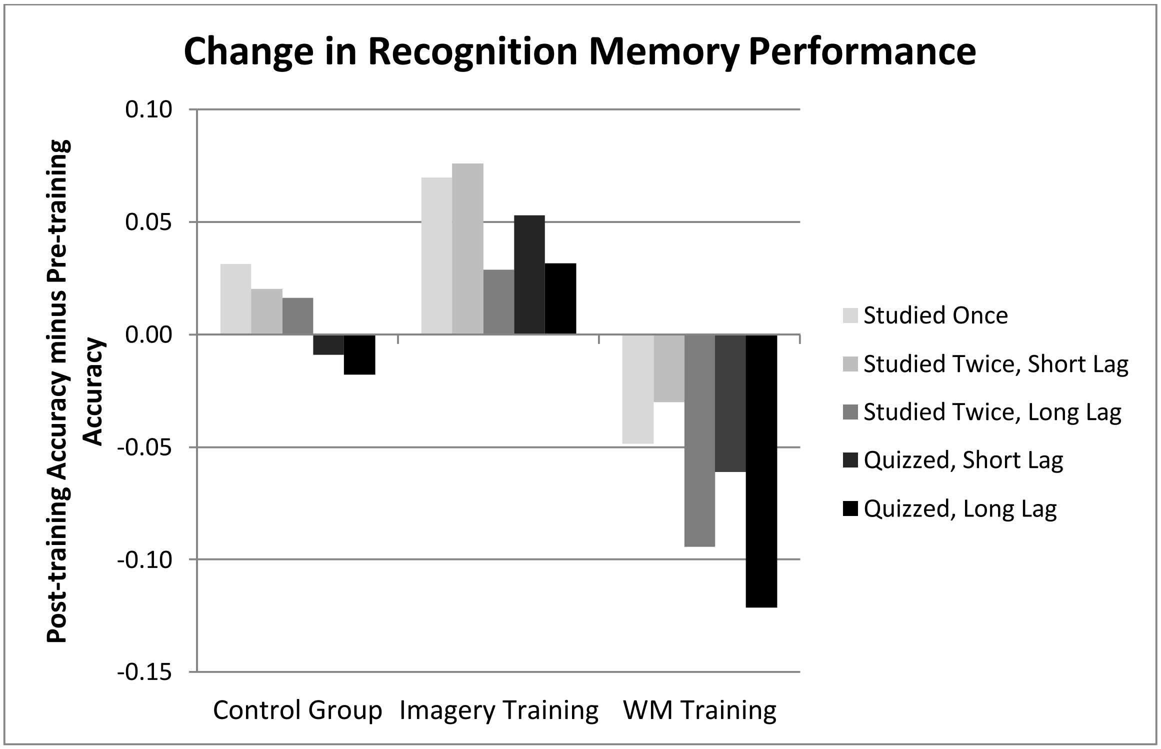

Change in Recognition Memory Performance
Post-training Accuracy minus Pre-training Accuracy
0.10
0.05
0.00
-0.05
-0.10
-0.15
Control Group
Imagery Training
WM Training
Studied Once
Studied Twice, Short Lag
Studied Twice, Long Lag
Quizzed, Short Lag
Quizzed, Long Lag

Figure 5.

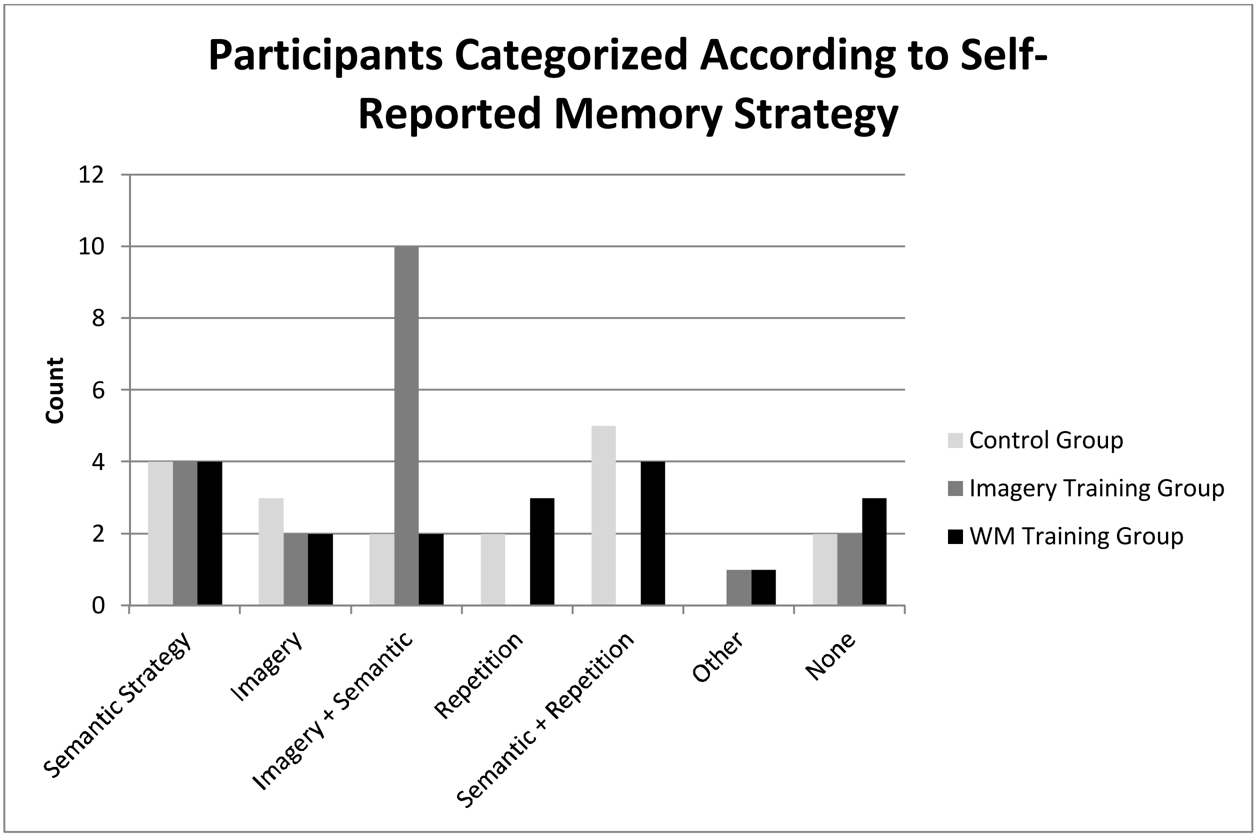
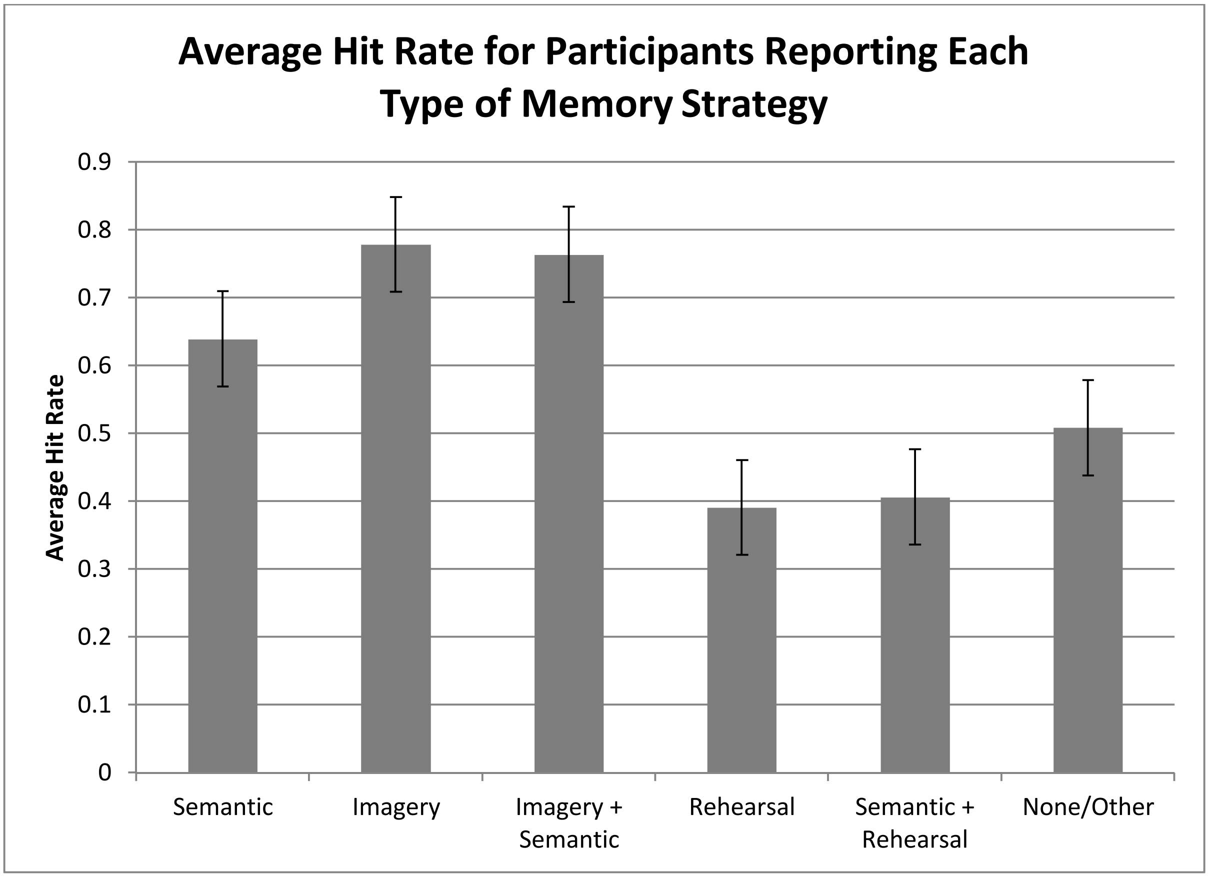

Participants Categorized According to Self-Reported Memory Strategy
Count
12
10
8
6
4
2
0
Semantic Strategy
Imagery
Imagery + Semantic
Repetition
Semantic + Repetition
Other
None
Control Group
Imagery Training Group
WM Training Group

Figure 6.


Average Hit Rate for Participants Reporting Each Type of Memory Strategy
Average Hit Rate
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Semantic
Imagery
Imagery + Semantic
Rehearsal
Semantic + Rehearsal
None/Other