Exploration of an adaptive training regimen that can target the secondary memory component of working memory capacity

**Bradley S. Gibson · Dawn M. Gondoli ·
William G. Kronenberger · Ann C. Johnson ·
Christine M. Steeger · Rebecca A. Morrissey**

Abstract The dual-component model contends that working memory (WM) capacity consists of two components. The first is a flexible attention component that involves the active maintenance of a limited amount of information in primary memory (PM), and the second is a controlled retrieval component that involves a cue-dependent search of secondary memory (SM) for information that has been lost from PM. Recent evidence has suggested that the adaptive WM training regimen known as "Cogmed-RM" is not optimally designed, because it only targets PM abilities, not SM abilities. The present study was conducted to investigate whether Cogmed-RM could be modified to target SM abilities by decreasing the recall accuracy threshold that defines individual ability during training. The main findings suggested that the SM component of WM capacity could be targeted by lowering the recall accuracy threshold. The present findings are important because they suggest that adaptive training regimens can be designed that selectively target specific components of WM capacity, and they raise the possibility that the potency of existing training regimens can be increased.

Keywords Working memory capacity · Working memory training · Dual-component model of working memory

B. S. Gibson (✉) · D. M. Gondoli · A. C. Johnson ·
C. M. Steeger · R. A. Morrissey
Department of Psychology, University of Notre Dame,
118 Haggar Hall,
Notre Dame, IN 46556, USA
e-mail: bgibson@nd.edu

W. G. Kronenberger
Indiana University School of Medicine, Indianapolis, IN, USA

Recent advances in neuroplasticity have raised the possibility that cognitive health may be optimized and preserved by engaging in training exercises that are specifically designed to target basic cognitive mechanisms. The societal implications of improved cognitive fitness are vast, and a recent market analysis suggested a growing public interest in these interventions, as expenditures increased from approximately $100 million in 2005 to approximately $225 million in 2007, with the largest increases occurring within the personal and healthcare segments of the market (Fernandez & Goldberg, 2008). However, despite the potential health benefits associated with cognitive-fitness regimens, empirical studies aimed at establishing the effectiveness of these interventions have generally lagged behind this growing public interest.

Critical to establishing the long-term utility of cognitive-fitness regimens is whether interventions can be designed that are flexible enough to maintain training effects outside the specific training environment—producing so-called *far-transfer effects*. Equally important is understanding the causal etiology of these far-transfer effects so that the mechanisms underlying cognitive enhancement can be ascertained; this goal has fostered the development of training regimens that are designed to train specific cognitive processes, rather than a complex mixture of different processes (Lustig & Flegal, 2008).

Some evidence now suggests that adaptive training of working memory (WM) can enhance higher-order cognitive abilities (see Buschkuehl & Jaeggi, 2010; Diamond & Lee, 2011; Klingberg, 2010; Melby-Lervåg & Hulme, 2012; Morrison & Chein, 2011; Shipstead, Hicks, & Engle, 2012; and Shipstead, Redick, & Engle, 2010, 2012, for recent reviews). For instance, recent empirical studies have been interpreted to suggest that training-induced increases in WM capacity can be accompanied by improvements in fluid intelligence (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Jausovec &

Jausovec, 2012; Klingberg, Fernell, Olesen, Johnson, Gustafsson, Dahlström and Westerberg 2005), reading comprehension (Chein & Morrison, 2010; Dahlin, 2010), math competence (Holmes, Gathercole, & Dunning, 2009), and attention-deficit hyperactivity disorder (ADHD) symptoms (Beck, Hanson, Puffenberger, Benninger, & Benninger, 2010; Gibson, Gondoli, Johnson, Steeger, & Morrissey, 2011; Holmes, Gathercole, Place, Dunning, Hilton, & Elliot 2010; Klingberg et al., 2005).

However, others have questioned the causal etiology of these effects by questioning whether the benefits of adaptive WM training are actually due to changes in WM capacity (Shipstead et al., 2012a, b). Hence, although adaptive WM training regimens may be capable of enhancing higher-level cognitive abilities, the causal etiology of these far-transfer effects remains poorly understood, and ironically may not include a role for WM capacity. There is thus a critical need to understand which components of WM capacity are targeted by existing training regimens.

Recently, Gibson and colleagues (Gibson et al., 2011; Gibson, Kronenberger, Gondoli, Johnson, Morrissey, & Steeger, 2012) have attempted to clarify the etiology of WM training within the context of the dual-component model of WM (Unsworth & Engle, 2007a; Unsworth & Spillers, 2010). According to this model, WM capacity is composed of at least two dissociable components: (1) the active maintenance of a limited amount of information in primary memory (PM) and (2) the retrieval of goal-relevant information from secondary memory (SM) after that information has been lost from PM (due to failures of active maintenance and/or storage limitations).

Gibson et al. (2011) investigated whether the PM or SM component of WM capacity, or both, could be enhanced by one well-known and widely used adaptive WM training regimen known as "Cogmed-RM," which contains a mixture of both verbal and spatial simple-span exercises. Because spatial simple-span tasks may engage the components of WM capacity more than verbal simple-span tasks (Kane, Hambrick, Tulholski, Wilhelm, Payne, & Engle, 2004; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Oberauer, 2005; Shah & Miyake, 1996), the exercises were divided into two separate training conditions—a verbal training condition ($N = 20$) and a spatial training condition ($N = 17$)—to examine whether spatial training might engage the SM component more than verbal training.

Following Unsworth and Engle (2007a), the numbers of items recalled from PM and SM, as well as recall accuracy as a function of serial position, were obtained from performance on verbal and spatial immediate free recall tasks. The main findings showed that Cogmed-RM selectively improved the number of items recalled from PM ($d = 0.52$), but not the number of items recalled from SM ($d = 0.15$). Consistent with this interpretation, a significant interaction

between serial position and time was also observed when recall accuracy was analyzed, indicating that improvement was confined solely to the recency portion of the serial-position curve. Furthermore, the same pattern was observed across both the verbal and spatial training conditions.

Gibson et al.'s (2011) findings have important practical implications for the design of WM training regimens, because other studies have suggested that the ability to retrieve information from SM is just as important as the ability to actively maintain information in PM, if not more so, for explaining individual differences in WM capacity (Unsworth & Engle, 2007a; Unsworth & Spillers, 2010; Unsworth, Spillers, & Brewer, 2010), fluid intelligence (Mogle, Lovett, Stawski, & Sliwinski, 2008; Unsworth, Brewer, & Spillers, 2009; Unsworth & Engle, 2007b; Unsworth & Spillers, 2010; Unsworth et al., 2010), and ADHD symptoms (Gibson, Gondoli, Flies, Dobrzenski, & Unsworth, 2010). On the basis of these findings, there is good reason to believe that the most potent WM training regimens would be those that can target both PM and SM abilities. Hence, the potential benefits of WM training regimens such as Cogmed-RM are not as potent as they could be.

In addition, Gibson et al.'s (2011) findings may also shed light on Shipstead et al. (2012a) conclusion that WM training rarely has been shown to enhance the capacity of WM, as measured by complex-span tasks. Complex-span tasks require dual-task performance and may provide better measures of SM than of PM abilities, because the processing task causes all but the last of the to-be-remembered list items to be displaced from PM into SM (Unsworth & Engle, 2007b; see also Chein, Moore, & Conway, 2011). As a result, successful recall in complex span tasks mostly reflects the retrieval of information from SM.

In contrast, simple-span tasks may provide better measures of PM than of SM abilities because the displacement of items from PM into SM only occurs with relatively long list lengths in these tasks (i.e., with list lengths that exceed the storage capacity of PM; Unsworth & Engle, 2007b). As a result, successful recall in simple-span tasks mostly reflects the retrieval of information that is actively maintained in PM, at least when the list length is relatively short. However, successful recall in simple-span tasks may increasingly measure SM abilities (as opposed to PM abilities) as list length increases beyond the storage capacity of PM (see also Unsworth & Engle, 2006). Given that Cogmed-RM appears to target PM more than SM abilities, it is perhaps not too surprising that such training does not consistently enhance performance on complex-span tasks.

With these considerations in mind, Gibson et al. (2012) recently investigated whether Cogmed-RM could be modified to target the SM component by converting its standard simple-span exercises into complex-span exercises (see also Chein & Morrison, 2010). This modification

was accomplished by inserting additional processing tasks between to-be-remembered list items in a critical subset of both verbal and spatial exercises, similar to the operation span (Turner & Engle, 1989) and symmetry span tasks (Kane et al., 2004), respectively, and two separate training conditions were compared: a standard-exercise training condition ($N = 31$) and a modified-exercise training condition ($N = 30$).

If inserting an additional processing task causes all but the last of the to-be-remembered list items to be displaced from PM into SM, then training with adaptive complex-span exercises should target SM abilities more than does training with adaptive simple-span exercises. Thus, SM abilities might be enhanced to a greater extent following training in the modified-exercise condition than in the standard-exercise condition. For this reason, the modified-exercise condition was construed as the treatment condition in Gibson et al.'s (2012) study, whereas the standard-exercise condition was construed as the control condition.

Using the same outcome measures as Gibson et al. (2011), Gibson et al. (2012) found that the standard-exercise training condition selectively improved the number of items recalled from PM ($d = 0.36$), but not the number of items recalled from SM ($d = 0.04$). As such, these findings corroborated the findings reported by Gibson et al. (2011). However, despite evidence that the complex-span exercises were more distracting than the simple-span exercises across the entire duration of the training period, the same pattern of results was also observed in the modified-exercise training condition: Namely, the number of items recalled from PM was improved ($d = 0.47$), but the number of items recalled from SM was not ($d = 0.03$). In addition, a significant interaction between serial position and time was also observed in both training conditions when recall accuracy was analyzed, indicating that improvement was confined solely to the recency portion of the serial-position curve. On the basis of these findings, Gibson et al. (2012) concluded that converting simple-span exercises into complex-span exercises is not sufficient to target the SM component of WM capacity, perhaps because the insertion of the processing task does not always cause to-be-remembered list items to be displaced from PM.

Although the use of complex-span tasks may increase the probability that any given item is lost from PM during training, satisfaction of this criterion alone does not guarantee that trainees will be given adequate opportunities to practice retrieving this information from SM. Rather, providing adequate opportunities to practice retrieving information from SM may require further consideration of how the span length of the adaptive exercises is adjusted on a trial-by-trial basis to match the WM capacity of the trainee.

There are at least two reasons to suspect that the adaptive algorithm used in the standard version of Cogmed-RM is biased to target PM but not SM abilities. First, the recall accuracy threshold used to adjust list length in standard

versions of Cogmed-RM (Gibson et al., 2011; Holmes et al., 2009; Klingberg et al., 2005), as in other span-based adaptive training regimens (Chein & Morrison, 2010), has been universally set at 100 %. As a result, the length of the upcoming list will not increase until the trainee can consistently recall all the items on the current list with perfect accuracy. Second, recall from SM tends to be less accurate than recall from PM (Unsworth & Engle, 2007a). This is because recall from SM involves a probabilistic search through a representation of both relevant and irrelevant items, whereas recall from PM has been construed as simply "unloading" the contents of PM (Unsworth, 2007; Unsworth et al., 2009; Unsworth & Engle, 2007a).

If recall from SM is harder and less accurate than recall from PM, then the use of a 100 % recall accuracy threshold may constrain full engagement of the SM component. For instance, consider an individual who is training with a 100 % recall accuracy threshold, and consider that this individual has just encountered a list that exceeded the storage capacity of PM by one item. Let us suppose further that this individual was able to recall all of the items that were being maintained in PM with perfect accuracy, but failed to recall the one item that was lost from PM and had to be retrieved from SM. Because list length is contingent on perfect recall in this context, the length of the next list will be decreased by one item. In this way, a 100 % recall accuracy threshold may enable this individual to train at the maximal (or near maximal) storage capacity of PM, without providing much opportunity to train retrieval from SM.

In contrast, now suppose that this individual had been training with a lower recall accuracy threshold. Although, recall failed for the one item that was lost from PM, the length of the next list would not decrease, but rather would continue to increase until this individual was unable to satisfy the lower recall accuracy threshold. Consequently, this individual would now be given more opportunity to practice retrieving list items from SM, and as a result, his or her ability to retrieve might improve, and SM ability could increase.

In summary, increased engagement of the SM component during training may require decreasing the recall accuracy threshold from 100 % to a lower value. A decrease in the recall accuracy threshold will likely elicit more retrieval from SM before recall is terminated on any given training trial, and it will also ensure that list length is determined more by the limitations of SM abilities and less by the limitations of PM abilities.

The present study

Researchers who attempt to develop novel, theoretically inspired WM training regimens should not be expected to proceed directly from abstract theory to costly large-scale randomized controlled trials. Rather, successful development typically requires one or more exploratory studies to ensure

that the training regimen is operating as intended (Leon, Davis, & Kraemer, 2011). Indeed, our two previous training studies (Gibson et al., 2011, 2012) failed to find any significant change in the SM component over time. Consequently, a more exploratory analytic strategy was viewed as a necessary (and more feasible) first step in the present study.

Accordingly, the primary purpose of the present study was to explore whether using a lower recall accuracy threshold during training could influence SM abilities. If lowering the recall accuracy threshold can target the SM component, then significant enhancement of SM abilities should be observed across time in the present study. Furthermore, significant enhancement of PM abilities should also be observed across time in the present study, regardless of whether significant enhancement of SM abilities was observed.

In addition to examining whether using a lower recall accuracy threshold during training can target the SM component, in the present study we also examined whether the magnitude of this effect might interact with exercise type. Accordingly, the lower recall accuracy threshold was implemented within both the standard-exercise (simple-span) and modified-exercise (complex-span) training conditions used by Gibson et al. (2012). According to Unsworth and Engle (2007b), inserting an additional processing task between to-be-remembered items (as in complex-span tasks) should cause distraction and increase the probability that list items are lost from PM, regardless of list length. If so, then the average span achieved during training in the modified-exercise condition should be consistently lower than the average span achieved during training in the standard-exercise condition (Gibson et al., 2012). Furthermore, if lowering the recall accuracy threshold interacts with exercise type, then greater enhancement of SM abilities may be observed across time in the modified-exercise training condition than in the standard-exercise training condition.

If significant change in the SM component of WM capacity were to be observed over time in the present study, then it would be reasonable to progress to the second stage of analysis, which would explore whether the observed patterns of enhancement could be distinguished from a control condition. This analysis would compare the active training conditions to a no-contact control condition that was not expected to enhance either component of WM capacity. This comparison would enable us to determine whether the observed patterns of enhancement could be distinguished from test–retest effects.

Method

Participants

A total of 20 undergraduates from the University of Notre Dame were recruited and randomly assigned to either the standard-exercise ($N = 10$) or the modified-exercise ($N = 10$) training condition. Each participant was paid a total of $100.00 for their participation (pretraining assessment, five-week intervention, and posttraining assessment).

Pretraining and posttraining assessments

Consistent with previous studies (Gibson et al., 2011, 2012), verbal and spatial immediate free recall (IFR) tasks were used to measure the PM and SM components of WM capacity in the present study. According to Unsworth and Engle (2007a), IFR tasks are valid measures of WM capacity. In their reanalysis of Engle, Tuholski, Laughlin, and Conway's (1999) structural equation model, Unsworth and Engle (2007a) showed that performance on a verbal IFR task loaded just as highly on the latent construct of WM capacity as did performance on three more traditional complex-span tasks: IFR (.73), operation span (.77), reading span (.58), and counting span (.62). Unsworth et al. (2010) replicated these results with a different sample, and they further reported that the split-half reliability of this verbal IFR task is .85.

In the present study, one verbal and one spatial IFR task was administered immediately before and within one week of finishing the intervention (Gibson et al., 2010, 2011, 2012). In these tasks, 15 lists of 12 unique high-frequency words or spatial locations were presented. The spatial locations were randomly selected from a 15×15 matrix. Each item was presented consecutively for 1 s. Following the presentation of a single list, question marks appeared in the center of the screen, prompting a response by the participant. Participants were given 30 s to recall as many of the words or spatial locations from the current trial as possible, in any order that they wished. Words were reported orally and recorded digitally, whereas spatial locations were reported by clicking a mouse at the appropriate locations and stored by the computer. Following previous research (Gibson et al., 2011, 2012), participants were explicitly instructed to begin recalling words or spatial locations toward the end of the list to control for recall initiation strategies, though strict serial ordering was not required (see also Craik & Birtwistle, 1971). For each task, three practice trials preceded the experimental trials.

According to Unsworth and Engle (2007a), IFR tasks may be better suited for assessing recall from PM and SM than are complex- or simple-span tasks, because IFR tasks can provide separate measures of each component, whereas complex- and simple-span tasks typically provide a single measure that may reflect contributions from both components. For instance, Tulving and Colotla (1970) developed a method that can be applied to free recall that estimates the numbers of items recalled from PM and SM (Gibson et al., 2010, 2011, 2012; Unsworth & Engle, 2007a; Unsworth et al., 2010).

According to Tulving and Colotla (1970), estimates of the number of items that can be recalled from PM and SM must take into consideration both input and output interference; the greater the amount of interference preceding recall of an item, the more likely will the item be recalled from SM as opposed to PM. Following Tulving and Colotla, the number of items between a given item's presentation and its recall is tallied. An item is considered to be recalled from PM when there are seven or fewer items intervening between that item's presentation and its recall. In contrast, an item is considered to be recalled from SM when more than seven items intervene between that item's presentation and its recall. Other researchers (Craik & Birtwistle, 1971; Unsworth & Engle, 2007a) have validated these estimates by showing that recall from SM was affected by the buildup of proactive interference, whereas recall from PM was not (see also Watkins, 1974).[1]

One concern with using Tulving and Colotla's (1970) method to estimate the PM and SM components of WM capacity within the present context is that this method uses relatively coarse and rigid criteria for defining PM and SM, which may not be optimal for measuring change in these components. For instance, selective enhancement of the PM component following training could be misattributed to the SM component if improvement in the number of items recalled from PM expanded beyond the fixed criterion for PM. As a result, one must be vigilant that this method does not underestimate the number of items recalled from PM and overestimate the number of items recalled from SM.

One way to address this concern would be to also analyze recall accuracy on the two IFR tasks as a function of serial position (see also Gibson et al., 2011, 2012). For instance, Unsworth and Engle (2007a) compared the performance of high- and low-capacity individuals on a 12-item verbal IFR task and found significant differences in recall accuracy across all but the last serial position. This finding led them to conclude that high-capacity individuals differ from low-capacity individuals in terms of both PM and SM abilities. These findings suggest that it ought to be possible to improve performance across most (if not all) serial positions if WM training is truly able to increase both PM and SM abilities. In contrast, if the apparent improvement in SM abilities actually reflects improvement in PM abilities, then such improvement should be confined to the recency portion

of the serial-position curve when recall accuracy is analyzed (see also Gibson et al., 2011, 2012).

WM training interventions

The participants in both training conditions were instructed to complete 25 days of WM training within five weeks. Following the previous protocol (see Gibson et al., 2011, 2012), participants were required to complete at least 20 days of WM training within this five-week period to be included in the final analyses. The participants completed the computerized WM training at home via the Internet. Daily training performance (maximum span, minimum span, and average span) on each exercise was logged to a secure website and monitored on a daily basis to ensure compliance. Both training conditions included a combination of verbal and spatial span exercises (see Holmes et al., 2009, for a more detailed description of these exercises). The verbal exercises involved remembering the correct forward serial order of letters and digits, whereas the spatial exercises involved remembering the correct forward serial order of locations in a two- or three-dimensional grid. Each item was presented for 1 s. Note that a new list of items was presented on each trial. Note also that only eight of the ten possible exercises were presented on each day. Trainees completed all eight exercises each day; the total time spent training each day was set at 30 min (not including breaks).

Four of the exercises that were presented each day were designated as "common exercises," and they were selected from a total set of six exercises. In this way, trainees were introduced to different common exercises throughout the course of the training (every 5 days) to break the monotony. The common exercises were simple-span tasks, and the six common exercises used in the standard-exercise training condition were identical to those used in the modified-exercise training condition. These exercises were included to provide a common basis for comparison across the two training conditions.

The remaining four exercises (two verbal and two spatial) were designated as "critical exercises," and these same four exercises were presented every day. The four critical exercises used in the standard-exercise training condition were identical to those used in the modified-exercise training condition, except that the exercises used in the standard condition were simple-span tasks, whereas those used in the modified condition were complex-span tasks. The two critical verbal exercises were converted to complex-span tasks by inserting basic mathematical operations [e.g., $(2 + 2)/4 = 3$] between list items (also digits), as in the operation span task. These operations were considered to be of intermediate and optimal difficulty by Turner and Engle (1989). Both the interim and final solutions to the operation were always a whole number between 0 and 9. This resulted in a

---

[1] Unsworth et al. (2010) reported the split-half reliabilities of these measures of PM and SM to be .62 and .76, respectively. However, there is good reason to believe that Unsworth et al. (2010) underestimated the reliability of these measures, because they did not control the recall initiation strategy used in their IFR task. This is a potential problem because the recall initiation strategy can change across trials, and this change can influence the measures of PM and SM (Unsworth, Brewer, & Spillers, 2011). For this reason, the recall initiation strategy was controlled in the present study.

total pool of 51 operations that were paired equally with correct and incorrect final solutions. Participants responded whether the operation was "true" or "false" before the next list item was presented.

Likewise, the two critical spatial tasks were converted to complex-span tasks by inserting random-dot spatial patterns between list items (spatial locations), as in the symmetry span task (Kane et al., 2004). These spatial patterns were created by randomly filling half of the cells in an 8 × 4 matrix. This resulted in a total pool of approximately 600 billion different spatial patterns; these patterns were either repeated or not in an identical grid to form vertically symmetrical or asymmetrical patterns. When the patterns were asymmetrical, the pattern on one side differed by two dots relative to the pattern on the other side. Participants responded whether the pattern was "symmetrical" or "asymmetrical" before the next list item was presented.

Participants were required to maintain 100 % accuracy on the processing tasks to ensure that they did not ignore these tasks, and no time limit was imposed on the performance of the processing task. Failure to maintain 100 % accuracy on the processing tasks for any given trial nullified the recall performance for that trial, and the trial was considered unsuccessful. The next list item was presented immediately following each response on the processing tasks. Note that the two processing tasks used in the modified-exercise condition (mathematical operations and symmetry) were also included as *separate* exercises (not interleaved between list items) in the standard-exercise condition, to control for training time and simple exposure to this information.

Each training exercise always began with two-item lists on the first day of training. The number of list items presented on each subsequent trial was adjusted automatically, on a trial-by-trial basis, to match the WM span of the participant on each task, and the same adaptive algorithm was used in both training conditions. The recall accuracy threshold was modified in the present study, such that it decreased as list length increased, resulting in an overall recall accuracy threshold that was less than 100 %. More specifically, no errors were allowed during recall on any given trial when training spans ranged from two to four items; one error was allowed during recall on any given trial when training spans ranged from five to seven items; two errors were allowed during recall on any given trial when training spans ranged from eight to ten items; three errors were allowed during recall on any given trials when training spans ranged from 11 to 13 items; and so on. For instance, the length of the training span increased from five to six items if participants achieved 100 % on the processing tasks and successfully recalled the correct forward serial order of 80 % of the five-item lists three times in row; in contrast, the length of the training span decreased from five to four items if participants failed to achieve 100 % on the processing

tasks or failed to recall the correct forward serial order of 80 % of the five-item lists three times in a row.

Results and discussion

Of the 20 participants who completed the pretraining assessment, all 20 continued on to the WM training phase of the study. Of these 20 participants, only one participant failed to complete at least 20 days of training, in the standard-exercise training condition. A total of 19 participants completed the posttraining assessments.

WM training: Critical exercises

The average spans achieved on the critical exercises are shown in Fig. 1 as a function of exercise modality and training duration in each of the two training conditions. Examination of the average spans achieved by each participant on the critical exercises revealed one participant in the modified-exercise training condition whose training trajectory far exceeded the trajectories of the other 18 participants in either training condition (see Fig. 1). Because this one participant had an inordinate effect on the overall performance of the modified-exercise training condition, we therefore excluded this participant from the remaining analyses.

The average spans achieved on the critical exercises were analyzed using a three-way, mixed analysis of variance (ANOVA), with Exercise Modality (verbal vs. spatial) and
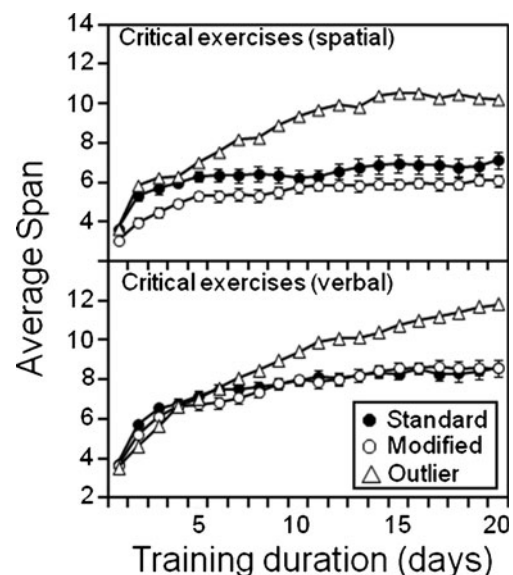

Fig. 1 Average spans achieved on the critical exercises, depicted as a function of training duration and training condition in each of the spatial (top panel) and verbal (bottom panel) exercise conditions. The one participant in the modified-exercise training condition whose training trajectory differed from the others is also plotted, for comparison. Error bars represent standard errors of the means

Training Duration (Days 6–20) as the two within-subjects factors, and Training Condition (standard-exercise vs. modified-exercise) as the sole between-subjects factor. Because the training spans always began at the same level for each individual, regardless of training condition, there was less chance for variation to occur during the early days of training. For this reason, Training Days 1–5 were excluded from the present analyses. However, the same results were obtained regardless of whether or not these early training days were included in the analysis.

As expected, average span length increased over time in both training conditions, as indicated by a significant main effect of training duration, $F(14, 224) = 20.33$, $p < .0001$, $\eta_p^2 = .56$. In addition, lower average spans were achieved when spatial exercises were performed ($M = 6.20$ items) than when verbal exercises were performed ($M = 8.08$ items), as indicated by a significant main effect of exercise modality, $F(1, 16) = 299.41$, $p < .0001$, $\eta_p^2 = .95$. More importantly, the main effect of training condition did not approach significance in this experiment, $F(1, 16) = 1.55$, $p > .20$, $\eta_p^2 = .09$. However, there was a significant interaction between exercise modality and training condition, $F(1, 17) = 5.31$, $p < .05$, $\eta_p^2 = .24$. As expected, subsequent analyses revealed that the average spans achieved in the modified-exercise training condition ($M = 5.77$ items) were significantly lower than those achieved in the standard-exercise training condition ($M = 6.64$ items) when the spatial exercises were compared, $t(16) = 2.37$, $p < .05$. In contrast, no difference was observed between the modified-exercise and standard-exercise training conditions (Ms = 8.06 and 8.09 items, respectively) when the verbal exercises were compared, $t(16) = 0.07$, $p > .90$. In addition, further evidence that average spans increased as a function of training duration (Days 6–20) was provided by regression analyses that revealed significant positive slopes in each of the four Exercise Modality × Training conditions (ps ranged from .006 to .0001; betas ranged from .237 to .495; R^2 values ranged from .056 to .245).

Despite using identical processing tasks, the pattern of average spans observed across the two training conditions in the present study differed somewhat from the pattern of average spans observed by Gibson et al. (2012). Using a heterogeneous sample of adolescents (9–16 years of age), the previous study reported a significant main effect of training condition, indicating that both the operation and symmetry judgment tasks consistently caused distraction. However, the results obtained in the present study suggested that only the symmetry judgment task consistently caused distraction.

Of course, the pool of processing items was several orders of magnitude smaller in the operation judgment task than in the symmetry judgment task, which raises the possibility that the operation judgments might have become less

distracting as participants became more familiar with these items. Indeed, although we did not include the first 5 days of training in our analysis of average spans reported above, these spans were found to be significantly lower in the modified-exercise training condition than in the standard-exercise training condition on the second day of training (Ms = 5.21 vs. 5.71 items, $p < .02$), and marginally lower on the third day of training (Ms = 6.10 vs. 6.54 items, $p = .08$), but not on subsequent days. Thus, a more refined conclusion is that the operation judgments did cause significant distraction in the present study, but only temporarily. Moreover, these same operation judgments may have caused more persistent distraction in Gibson et al.'s (2012) study because it is likely that the younger sample of participants used in that study were less familiar with the results of these operations (see also the supplementary materials for a detailed analysis of the mean correct processing latencies associated with the operation and symmetry judgment tasks).

WM training: Common exercises

Analysis of the average spans achieved on the critical exercises suggested that the participants in the modified-exercise training condition were exposed to significant distraction on at least some of the training exercises (the critical spatial exercises). We can get an initial impression of the impact of this more-difficult training by examining potential group differences in the average spans achieved over time on the common exercises. The average spans achieved on the common exercises are shown in Fig. 2 as a function of training duration in each of the two training conditions. As can be seen in Fig. 2, average spans were equal across the two training conditions, indicating that participants in the modified-exercise training condition did not develop higher WM capacity than participants in the standard-exercise training condition. More specifically, we found a significant main effect of
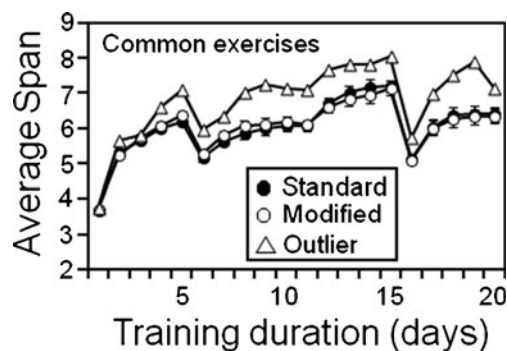

Fig. 2 Average spans achieved on the common exercises, depicted as a function of training duration in each of the two training groups. The one participant in the modified-exercise training condition whose training trajectory differed from the others is also plotted, for comparison. Note that some new exercises were introduced on Days 6, 11, 16, and 21 (not shown). Error bars represent standard errors of the means

training duration, $F(14, 224) = 74.08$, $p < .0001$, $\eta_p^2 = .82$, but neither the main effect of training condition nor the Training Duration × Training Condition interaction approached significance (both Fs ≤ 1). Although examination of the average spans achieved on the common exercises did not reveal any differences between the two training conditions over time, this common measure did not differentiate between the two components of WM capacity. Accordingly, we turn next to an analysis of our main WM outcome measures.

Stage 1 analysis: PM and SM

The analysis of WM outcomes proceeded in two stages, and the second stage was contingent on the findings obtained in the first stage. The first stage of analysis explored two primary questions: The first concerned whether using a reduced recall accuracy threshold during training could target the SM component, and the second concerned whether the beneficial effects of this reduced recall accuracy threshold might increase even further when paired with complex-span exercises (as in the modified-exercise training condition).

A three-way, mixed ANOVA was performed on the numbers of items recalled, with Time (pretraining vs. posttraining) and Memory Type (PM vs. SM, as derived using Tulving and Colotla's, 1970, method) as the two within-subjects factors and Training Condition (standard-exercise vs. modified-exercise) as the sole between-subjects factor. For the sake of simplicity, task modality (verbal IFR task vs. spatial IFR task) was not included as an independent variable, because a preliminary analysis indicated that task modality did not interact with any of the other experimental variables. In addition, although the results reported below excluded the participant with the superior training performance, the same results were obtained regardless of whether or not this participant was included.

The number of items correctly recalled from PM and SM is shown in Table 1 as a function of time for both the standard-exercise and modified-exercise training conditions. We found a significant main effect of time, $F(1, 16) = 19.21$, $p < .0001$, $\eta_p^2 = .55$, indicating that more items were recalled in the posttraining condition ($M = 3.35$) than in the pretraining condition ($M = 2.92$). In addition, a significant main effect of memory type also emerged, $F(1, 16) = 15.05$, $p < .001$, $\eta_p^2 = .48$, indicating that more items were recalled from PM ($M = 3.44$) than from SM ($M = 2.83$). However, the main effect of training condition did not approach significance, $F < 1$. Furthermore, although the effect of time (posttraining scores − pretraining scores) resulted in a numerically larger increase in the number of items that could be recalled from SM ($M = 0.60$ items), $t(17) = 4.02$, $p < .001$, than from PM ($M = 0.26$ items), $t(17) = 3.31$, $p < .005$, the Time × Memory Type interaction did not attain significance in

Table 1 Mean estimates the numbers of items recalled from secondary memory (SM) and primary memory (PM) as a function of time and training condition

	Time	
	Pretraining	Posttraining
Standard Training		
SM abilities	2.52 (0.26)	3.12 (0.19)
PM abilities	3.26 (0.11)	3.54 (0.12)
Modified Training		
SM abilities	2.55 (0.21)	3.15 (0.26)
PM abilities	3.36 (0.10)	3.60 (0.09)
No Contact		
SM abilities	2.52 (0.16)	2.54 (0.18)
PM abilities	3.29 (0.08)	3.37 (0.09)

Standard errors appear in parentheses

this study, $F(1, 16) = 2.81$, $p > .10$, $\eta_p^2 = .15$. None of the other interactions approached significance (all Fs < 1).

Stage 1 analysis: Serial-position effects

Might the significant enhancement of the SM component reported above be due to enhanced recall from PM that had spilled over into the SM range of measurement? If so, then corresponding improvements in recall accuracy should be confined to the recency portion of the serial-position curve, which in turn should result in a significant Time × Serial Position interaction. The probabilities of correct recall are shown in Fig. 3 as a function of time and serial position for both the standard-exercise (top panel) and modified-exercise (middle panel) training conditions.

A three-way, mixed ANOVA was performed on the probabilities of correct recall, with Time and Serial Position (1 to 12) as the two within-subjects factors, and Training Condition as the sole between-subjects factor. A significant main effect of time was apparent, $F(1, 16) = 19.21$, $p < .0001$, $\eta_p^2 = .55$, indicating that accuracy was higher in the posttraining than in the pretraining condition. We also found a significant main effect of serial position, $F(11, 176) = 130.26$, $p < .0001$, $\eta_p^2 = .89$, indicating that accuracy was higher for the recency items than for the prerecency items. However, neither the main effect of training condition nor any of the two- or three-way interactions approached significance in this study (all Fs < 1).

Thus, recall accuracy appeared to be consistently improved across the entire range of serial positions. This finding is consistent with Unsworth and Engle's (2007a) finding that high-capacity individuals differed from low-capacity individuals across the entire range of serial positions. Furthermore, this finding suggests that the significant enhancement of the SM component observed following
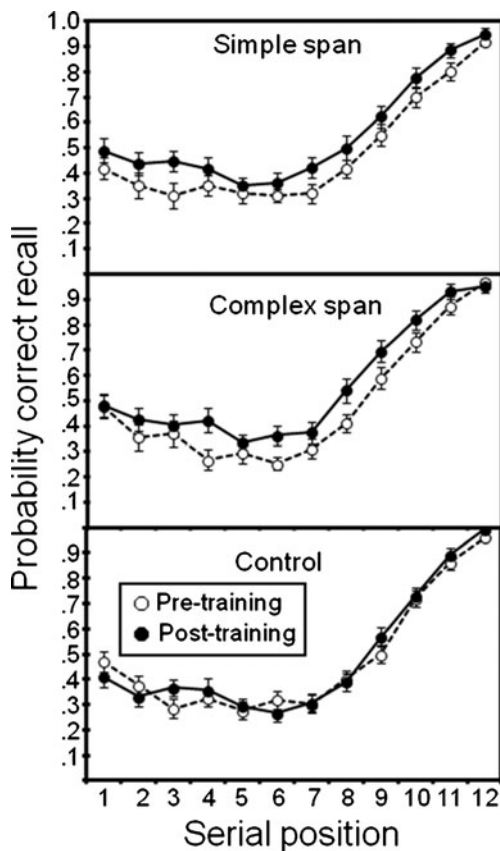
Fig. 3 Proportions of correct recall, depicted as a function of time and serial position in each of the standard-exercise (top panel), modified-exercise (middle panel), and no-contact control (bottom panel) conditions. Error bars represent standard errors of the means

training was not simply an artifact of how the PM and SM components were measured in the present study.

In summary, the first stage of our analyses provided evidence that the SM component can be targeted and enhanced by using a reduced recall accuracy threshold. However, the beneficial effects of this reduced recall accuracy threshold did not depend on training condition.

Stage 2 analyses: Comparison to a no-contact control condition

Having successfully demonstrated that reducing the recall accuracy threshold can target and enhance the SM component (as well as the PM component), the study progressed to the second stage of analysis, which attempted to distinguish performance in the two active training conditions from a no-contact control condition, in order to rule out simple test–retest effects. Accordingly, a new sample of 12 undergraduates was recruited from the University of Notre Dame approximately 1 year after the initial training study was completed. Despite this time gap, the participants in the control group completed the study at the same point in the academic year (spring semester) as the participants in the

two active training conditions, and they were also paid the same amount ($100.00) as the participants in the two active training conditions.

For the sake of comparison, the numbers of items recalled from PM and SM are listed in Table 1, and the probabilities of correct recall as a function of serial position are shown in Fig. 3 (bottom panel). As expected, preliminary analyses revealed no improvements in PM, SM, or recall accuracy as a function of time when the no-contact control condition was considered in isolation.

The present analyses therefore focused on a comparison of the numbers of items recalled from PM and SM between the two active training conditions (which were treated here as a single active treatment condition) and the no-contract control condition. Because pretraining estimates of the numbers of items recalled from PM and SM were significantly correlated with their corresponding posttraining estimates ($r = .635$ and $r = .420$, respectively; both ps < .02), we used an analysis of covariance (ANCOVA) approach in order to decrease error variance, and thereby increase the statistical power of our analysis (Maxwell & Delaney, 2004).

We began by conducting a one-way ANCOVA on posttraining estimates of the number of items recalled from SM, with Training Condition (active vs. control) as the between-subjects factor and pretraining estimates of the number of items recalled from SM as the covariate. As expected, we found a significant main effect of training condition, $F(1, 27) = 7.02$, $p < .01$, $\eta_p^2 = .21$, indicating that individuals recalled significantly more items from SM after training with a lower recall accuracy threshold (adjusted $M = 3.13$ items) than after no training at all (adjusted $M = 2.55$ items), after controlling for pretraining estimates of the number of items recalled from SM. This corresponds to a 23 % increase in the number of items recalled from SM.

In addition, we also conducted the same one-way ANCOVA on posttraining estimates of the number of items recalled from PM, with Training Condition as the between-subjects factor and pretraining estimates of the number of items recalled from PM capacity as the covariate. As expected, a significant main effect of training condition was found, $F(1, 28) = 4.59$, $p < .05$, $\eta_p^2 = .14$, indicating that individuals recalled significantly more items from PM after training (adjusted $M = 3.56$ items) than after no training at all (adjusted $M = 3.37$ items), after controlling for pretraining estimates of the number of items recalled from PM. This corresponds to a 6 % increase in the number of items recalled from PM. Altogether, these findings are important, because they suggest that the improvements in PM and SM capacity observed following active training were not due simply to practice with the WM outcome measures, as the individuals in this control condition were exposed to the same practice with these measures.

General discussion

In the present study, we investigated whether Cogmed-RM could be modified to target the SM component of WM capacity. Two modifications were investigated. The first involved decreasing the recall accuracy threshold in order to accommodate the more difficult task of retrieving information from SM, and the second involved converting the standard simple-span exercises into complex-span exercises in order to increase the likelihood that information might be lost from PM. In addition, the present study was also conducted within an exploratory context to investigate whether the potentially beneficial effects of either or both of these two theoretically inspired modifications could be observed before investing in a full-scale randomized controlled trial study (RCT).

With respect to the first modification, the present study provided important new evidence that the SM component can be targeted and enhanced when the recall accuracy threshold constraining training spans is reduced below 100 %. We have suggested that changes in recall accuracy threshold mainly affect the extent to which SM capacity is targeted during training, because this threshold is used to estimate individual abilities during training. Given that recall from SM is harder and less accurate than recall from PM (Unsworth & Engle, 2007a), the adaptive nature of the training regimen will become increasingly more likely to target the SM component as the recall accuracy threshold becomes less stringent (up to some point).

The algorithm used to lower the recall accuracy threshold in the present study was somewhat arbitrary and was chosen mainly because we thought that such values would create longer lists that individuals still had a reasonable chance of recalling. However, it is possible that algorithms that reduce recall accuracy thresholds even more may lead to greater enhancement of the SM component, owing to the fact that such lower thresholds should provide even greater opportunity to practice retrieving list items from SM.

With respect to the second modification, the present study showed that the SM component was enhanced equally, regardless of whether the reduced recall accuracy threshold was implemented within complex- or simple-span exercises. In addition, this finding was obtained despite the fact that at least some of the critical complex-span exercises appeared to operate as intended in the present study. In particular, the spatial complex-span exercises appeared to cause consistent distraction relative to the corresponding simple-span exercises, as evidenced by lower average spans achieved during training on the spatial complex-span exercises.

But if the distraction caused by the intervening symmetry judgments increased the likelihood that the spatial list items had to be retrieved from SM, why didn't training with complex-span exercises lead to greater enhancement of the SM component? Of course, one reason why this did not occur concerns the possibility that the complex-span exercises did not cause enough distraction in the present study, either because the pool of processing items was too small (at least in the verbal exercises) or because the duration of the processing period was not controlled.

However, a more likely reason that greater enhancement of the SM component was not observed in the modified-exercise relative to the standard-exercise condition concerns the relation between list length and SM involvement, on the one hand, and the use of an adaptive training algorithm to adjust list length, on the other. More specifically, the prediction that training with complex-span exercises should target the SM component more than training with simple-span exercises may only hold true when list length is held constant across the two training conditions. This is because, for a list of a particular length, list items are more likely to be retrieved from SM when the list is presented in the context of a complex-span task than when it is presented in the context of a simple-span task. As such, when list length is held constant, recall accuracy is likely to be lower in the complex-span than in the simple-span condition because recall from SM is more difficult and less accurate.

However, if the length of the simple-span list is allowed to grow longer than the length of the complex-span list, then the extents to which retrieval from SM is required for recall may become more equal across the two tasks (Unsworth & Engle, 2006). In fact, one way to equalize the extents to which retrieval from SM is required for recall across the two tasks would be to hold recall accuracy constant (and below 100 %) while allowing list length to vary—precisely the logic of contemporary adaptive WM training regimens. Thus, future investigations of exercise type will likely require nonadaptive training contexts in which list length is held constant while allowing recall accuracy to vary across training.

The present conclusion that the SM component of WM capacity can be targeted by reducing the recall accuracy threshold is important for at least two reasons. First, the present findings are important because they provide the necessary empirical foundation for proceeding to a full-scale RCT in which participants would be randomly assigned to either a decreased recall accuracy threshold condition, a 100 % recall accuracy threshold condition, or a placebo control condition. On the basis of both the present and previous findings (Gibson et al., 2012), these three training conditions can be interpreted to reflect a two-component (PM+ SM) training condition, a one-component (PM only) training condition, and a zero-component training condition, respectively (see the supplementary materials for preliminary evidence that the two-component condition can be empirically distinguished from the one-component condition).

Second, the present findings are also important because they provide the necessary empirical foundation for using WM training regimens as a testing ground for theories of WM capacity. For instance, according to the dual-component theory of WM capacity, individual differences in PM and SM abilities both explain unique variation in higher-level cognitive abilities such as fluid IQ. Consequently, a WM training regimen that can experimentally enhance both the PM and SM components of WM capacity should lead to greater enhancement of fluid IQ than would a WM training regimen that can experimentally enhance only the PM component of WM capacity. Thus, WM training represents an important tool for testing theories of WM capacity, just as theories of WM capacity represent an important tool for improving the potency of WM training.

Although the present study has provided preliminary evidence that reducing the recall accuracy threshold can target and enhance the SM component, future RCTs should also strive to provide a more thorough understanding of the nature of this enhancement. More specifically, current theories construe recall from SM as a multistep process. For instance, Unsworth (2007, 2009) has construed recall from SM in terms of three parameters: the size of the search set, the recovery of potential targets from this set, and error monitoring. Furthermore, using delayed recall and continuous distractor tasks that isolated the SM component from the PM component, Unsworth (2009) operationalized these three parameters in terms of recall latency, recall accuracy, and intrusion errors, respectively, in order to examine how preexisting individual differences in these three parameters related to preexisting differences in WM capacity and fluid IQ. The main findings suggested that preexisting differences in WM capacity and fluid IQ were primarily related to the use of smaller search sets (i.e., faster recall latencies) and better recovery of potential targets (i.e., higher recall accuracies) during retrieval.

Thus, future studies of WM training should attempt to clarify the nature of SM enhancement by using outcome tasks that can isolate the SM component and allow examination of the relative patterns of enhancement across recall latency, recall accuracy, and intrusion errors. Although the present findings indicating that SM abilities increased following training are commensurate with an increase in recall accuracy, the IFR tasks used in the present study cannot be used to provide pure measures of the size of the search set or error monitoring, because these tasks were used to measure both the PM and SM components of WM capacity.

In conclusion, the present study has provided a component analysis of WM training to examine whether the SM component of WM capacity could be targeted and enhanced by span-based exercises. The main findings suggested that the SM component could be enhanced by span-based exercises when a more lenient recall accuracy threshold was used. In contrast, the manipulation of exercise type (complex vs. simple span) showed little effect on the SM component of WM capacity thus far (see also Gibson et al., 2012). These findings are important because they raise the possibility that the effects of WM training on higher-level cognitive abilities such as fluid IQ can be increased by increasing the number of components that are targeted and enhanced by an intervention.

References

Beck, S. J., Hanson, C. A., Puffenberger, S. S., Benninger, K. L., & Benninger, W. B. (2010). A controlled trial of working memory training for children and adolescents with ADHD. *Journal of Clinical Child and Adolescent Psychology, 39,* 825–836.

Buschkuehl, M., & Jaeggi, S. M. (2010). Improving intelligence: A literature review. *Swiss Medical Weekly, 140,* 266–272.

Chein, J. M., Moore, A. B., & Conway, A. R. A. (2011). Domain-general mechanisms of complex working memory span. *Neuro-Image, 54,* 550–559.

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.

Craik, F. I. M., & Birtwistle, J. (1971). Proactive inhibition in free recall. *Journal of Experimental Psychology, 91,* 120–123.

Dahlin, K. I. E. (2010). Effects of working memory training on reading in children with special needs. *Reading and Writing, 24,* 479–491. doi:10.1007/s11145-010-9238-y

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333,* 959–964.

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. doi:10.1037/0096-3445.128.3.309

Fernandez, A., & Goldberg, E. (2008, March 11). *The state of the brain fitness/training software market 2008*. Retrieved from www.sharpbrains.com/blog/2008/03/11/report-the-state-of-the-brain-fitness-software-market-2008/

Gibson, B. S., Gondoli, D. M., Flies, A. C., Dobrzenski, B. A., & Unsworth, N. (2010). Application of the dual-component model of working memory to ADHD. *Child Neuropsychology, 16,* 60–79.

Gibson, B. S., Gondoli, D. M., Johnson, A. C., Steeger, C. M., Dobrzenski, B. A., & Morrissey, R. A. (2011). Component analysis of verbal versus spatial working memory training in adolescents with ADHD: A randomized, controlled trial. *Child Neuropsychology, 17,* 546–563.

Gibson, B. S., Kronenberger, W. G., Gondoli, D. M., Johnson, A. C., Morrissey, R. A., & Steeger, C. M. (2012). Component analysis of simple span vs. complex span adaptive working memory exercises: A randomized, controlled trial. *Journal of Applied Research in Memory and Cognition, 1,* 179–184.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12,* F9–F15.

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). Improved fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105,* 6829–6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108,* 10081–10086.

Jausovec, N., & Jausovec, K. (2012). Working memory training: Improving intelligence—changing brain activity. *Brain and Cognition, 79,* 96–106.

Kane, M. J., Hambrick, D. Z., Tulholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General, 133,* 189–217. doi:10.1037/0096-3445.133.2.189

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognition Sciences, 14,* 317–324.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD—A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 177–186.

Leon, A. C., Davis, L. L., & Kraemer, H. C. (2011). The role and interpretation of pilot studies in clinical research. *Journal of Psychiatric Research, 45,* 626–629.

Lustig, C., & Flegal, K. E. (2008). Targeting latent function: Encouraging effective encoding for successful memory training and transfer. *Psychology and Aging, 23,* 754–764.

Maxwell, S. E., & Delaney, H. D. (2004). Designing experiments and analyzing data: A model comparison perspective. Mahwah, NJ: Lawrence Erlbaum Associates.

Melby-Lervåg, M., & Hulme, C. (2012). Is working memory training effective? A meta-analytic review. *Developmental Psychology.* doi:10.1037/a0028228

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology. General, 130,* 621–640.

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological Science, 19,* 1071–1077.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18,* 46–60.

Oberauer, K. (2005). The measurement of working memory capacity. In O. Wilhelm & R. W. Engle (Eds.), *Understanding and measuring intelligence* (pp. 393–408). Thousand Oaks: Sage.

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. doi:10.1037/0096-3445.125.1.4

Shipstead, Z., Hicks, K. L., & Engle, R. W. (2012a). Cogmed working memory training: Does the evidence support the claims? *Journal of Applied Research in Memory and Cognition, 1,* 185–193.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50,* 245–276.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012b). Is working memory training effective? *Psychological Bulletin, 138,* 628–654. doi:10.1037/a0027473

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1,* 86–98.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154.

Unsworth, N. (2007). Individual differences in working memory capacity and episodic retrieval: Examining the dynamics of delayed and continuous distractor free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 1020–1034.

Unsworth, N. (2009). Variation in working memory capacity, fluid intelligence, and episodic recall: A latent variable examination of differences in the dynamics of free recall. *Memory & Cognition, 37,* 837–849.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There's more to the working memory–fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review, 16,* 931–937. doi:10.3758/PBR.16.5.931

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2011). Inter- and intra-individual variation in immediate free recall: An examination of serial position functions and recall initiation strategies. *Memory, 19,* 67–82.

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language, 54,* 68–80. doi:10.1016/j.jml.2005.06.003

Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104–132. doi:10.1037/0033-295X.114.1.104

Unsworth, N., & Engle, R. W. (2007b). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133,* 1038–1066.

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attentional control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language, 62,* 392–406.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2010). The contributions of primary and secondary memory to working memory capacity: An individual differences analysis of immediate free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 240–247. doi:10.1037/a0017739

Watkins, M. J. (1974). Concept and measurement of primary memory. *Psychological Bulletin, 81,* 695–711.