The prospects of cognitive (brain) training as an aid for teaching thinking

Oshin Vartanian, Erin L. Beatty

Published online on: 03 Jun 2015
Introduction

Cognitive (brain) training can be defined as targeted interventions meant to improve specific mental capacities or abilities through repeated practice. By and large, researchers have focused on working memory (WM) training. WM can be defined as “a multicomponent system for active maintenance of information in the face of ongoing processing and/or distraction” (Conway et al., 2005, p. 770). We now know that WM capacity is correlated positively with performance in a number of domains central to higher cognition including reasoning (De Neys, 2006), creative problem solving (De Dreu et al., 2012), and decision making (see Baddeley, 2003). Its ubiquity in human thought processes makes the current focus on training WM understandable. In essence, WM training offers the tantalizing prospect of improving a single domain-general capacity that can in turn contribute to better performance in multiple specific domains.

However, aside from the direct benefits of improving WM capacity for cognitive performance, there are at least two other indirect reasons for justifying the magnitude of the current thrust of research on WM training. First, data based on latent variable analysis show that WM capacity likely plays a causal role in fluid intelligence (Kane et al., 2004)—defined as the ability to adapt to new situations and to perceive new patterns and relationships (Cattell, 1963; Unsworth & Engle, 2005). Fluid intelligence is typically measured as a function of one’s ability to reason and solve problems in new contexts. In this sense, increasing WM capacity as a function of training would appear to have the potential to increase fluid intelligence. This transfer would be consequential because scores on psychometric measures of intelligence are known to predict numerous outcome measures of applied relevance such as school and job performance (see Neisser et al., 1996). According to this model, gains in fluid intelligence would mediate the link between improvements in WM performance and real-world outcome measures.
Second, impairments in WM capacity have been linked to some of the key behavioral indices of Attention-Deficit/Hyperactivity Disorder (ADHD), specifically inattention (Rapport, Chung, Shore & Isaacs, 2001). In turn, ADHD is associated with poor school performance and occupational outcomes (Barkley & Murphy, 2010; Biederman et al., 2012). Because WM is central to information processing in the face of distraction (Conway et al., 2005), there has been great interest in using WM training as a behavioral intervention tool to alleviate ADHD symptoms and associated functional deficits in school and job performance. In essence, demonstrating the efficacy of WM training, especially in classroom settings, would offer an alternative to medication which currently constitutes the most effective method for treating ADHD (Faraone & Buitelaar, 2010). In this way, reduction in ADHD symptoms would mediate the effect of WM training on school performance.

In summary, WM training can be used as a tool for aiding the teaching of thinking in at least three ways. First, directly, improvements in WM capacity should lead to improvement in performance on other tasks that draw on WM capacity. Second, indirectly, any gains in fluid intelligence that might accompany training-related gains in WM capacity should likely lead to improvements in tasks that draw on fluid intelligence. Third, and also via an indirect route, if gains in WM capacity were to lead to reductions in the symptoms that define ADHD, specifically distractibility, then one would expect training-related gains in WM capacity to improve school performance in children with ADHD. The organization of this chapter will reflect a discussion of our state of knowledge about each of these three research areas, and their potential contributions to improving cognitive performance.

Neuroplasticity and WM training

Importantly, although most of the evidence on WM training has been generated in behavioral studies, considerable work has also involved tracking training-related changes in brain function and structure. These studies are important for two reasons. First, they can be used to establish a mechanism whereby training-related improvements and transfer occur (Buschkuehl, Jaeggi & Jonides, 2012; Klingberg, 2010). Specifically, there is reason to believe that whether or not transfer occurs is a function of the extent to which the brain regions affected by WM training overlap with those recruited by the untrained tasks (see Dahlin et al., 2008). In turn, the greater the structural and functional similarities between the trained and untrained tasks (e.g., both necessitate the storage of information, suppression of distractors, etc.), the greater will be the likelihood that those sets of regions will overlap.

For example, we have recently shown that compared to an active control group that trained on a choice reaction time (RT) task, participants training on the n-back task performed significantly better on a delayed matching-to-sample task (dMTS) (Beatty et al., 2014). The n-back task requires the participant to decide, on a trial-by-trial basis, whether a stimulus presented in the current trial matches a target stimulus presented a specific number of trials earlier in the sequence. The letter n denotes the specific number of trials that separate the current trial from the target trial. In turn, dMTS is a WM task that necessitates the encoding, retention, and retrieval of stimulus representations in sequential order. Because we scanned participants in the functional magnetic resonance imaging (fMRI) scanner while they performed the dMTS, we were able to show that compared to the active control group, participants in the n-back group exhibited greater activation in the left inferior gyrus (IFG) only during the retention phase of dMTS, but not other phases (i.e., encoding and retrieval). This suggests that the mechanism of transfer from n-back training to dMTS likely involved maintenance, in the process recruiting a region (i.e., IFG) known to underlie delay-period maintenance in visual WM (Ranganath et al., 2004; de Zubicaray et al., 2001).

The second reason why tracking training-related neuroplasticity in the brain is important is that it can elucidate the direction of the transfer effect on the untrained task. This is especially
true if the transfer from WM training to the untrained task is hypothesized to be mediated by gains in fluid intelligence—typically assessed using the Raven’s *Advanced Progressive Matrices (RAPM)* (Raven et al., 1998). Interestingly, it has been shown that there are at least two different ways in which such a transfer effect could be realized in the prefrontal cortex (PFC)—a region frequently activated in WM training studies (Buschkuehl et al., 2012; Klingberg, 2010). On the one hand, there is reason to believe that improving fluid intelligence as a function of WM training should result in enhanced neural efficiency—typically operationalized as a lower blood oxygenation-level dependent (BOLD) signal—in the PFC. This is consistent with the inverse relation observed between fluid intelligence and metabolic rate in the parieto-frontal network using a variety of cognitive tasks that tap fluid intelligence (Deary et al., 2010; Jung and Haier, 2007; Neubauer et al., 2002, 2005; see also van der Heuvel et al., 2009), as well as indications that this inverse relation is most likely to be observed in the frontal cortex (Neubauer & Fink, 2009). Studies have shown that participants with higher scores on psychometric tests of intelligence—assessed using diverse measures of intelligence including the RAPM as well as instruments that provide subset scores in complete or abbreviated form (e.g., Wechsler, 1981)—recruit fewer regions in their brain for performing cognitively demanding tasks that draw on intelligence compared to participants with lower scores on psychometric tests of intelligence. On the other hand, there is also evidence for a positive correlation between neural activity and fluid intelligence (Gray et al., 2003; Lee et al., 2006; Luders et al., 2009), as well as data demonstrating an increase rather than a decrease in PFC activity following WM training (Olesen, Westerberg, & Klingberg, 2004). Establishing a direction for transfer effects is important because it has an impact on our interpretation of the change, either in the form of efficiency or greater recruitment. By accessing neural data, we therefore have a fuller understanding of the “joints in the system” that drive transfer effects.

**WM training: transfer to near and far tasks**

Recently, Melby-Lervåg and Hulme (2013) conducted a large-scale meta-analysis of 23 WM studies to assess their effectiveness on near and far transfer. In terms of the former, their results demonstrated that WM training was effective in producing improvements in WM performance. A number of the studies had also assessed WM performance at follow-up, ranging between three and eight months following the completion of training. Here, the analysis showed that whereas for verbal WM gains in performance were not sustained at follow-up, there was some evidence to suggest that gains in visuospatial WM were maintained at follow-up. Switching to far transfer, the results were less promising. They demonstrated that WM training did not confer any advantage to performance on other structurally dissimilar untrained tasks, including nonverbal and verbal ability, inhibitory processes in attention, word decoding, and arithmetic. The results of Melby-Lervåg and Hulme’s (2013) meta-analysis indicated that whereas WM training can be expected to improve performance on trained WM tasks, training-related gains do not transfer to other untrained tasks. Importantly, gains in visuospatial skills appear to be maintained better than gains in verbal skills, making them stronger candidates for reliable effects.

The results of Melby-Lervåg and Hulme’s (2013) meta-analysis dovetail with the results of a study conducted by Owen et al. (2010). This study is important for a number of reasons. First, of particular relevance for those researchers and practitioners interested in training cognition outside of standard laboratory settings, Owen et al. administered training to 11,430 participants online, who in turn trained several times each week for six weeks. Second, rather than limiting the intervention to WM training alone, the participants trained on a wide host of tasks designed to improve reasoning, memory, planning, visuospatial skills, and attention. The researchers observed improvements in every one of the cognitive tasks that were trained. However, no transfer effects were observed to untrained tasks, even in cases where the target
tasks were cognitively closely related to the trained tasks. A largely similar set of results were obtained in a more recent study in which a sample comprised of full-time and part-time staff from an Australian national public service organization engaged in cognitive training exercises across the domains of memory, attention, language, executive function, and visuospatial abilities—all delivered online (Borness et al., 2013). The outcome measures consisted of cognitive, psychological well-being, and productivity indices. Not only were no beneficial effects of cognitive training observed on any of the outcome measures, but the active control condition—comprised of viewing short documentaries about the natural world—experienced significant increases in their self-reported quality of life, decrease in stress levels, and overall improvement in psychological well-being assessed at six months follow up. Thus, it appears that if we widen our focus beyond WM training specifically, there does not appear to be much evidence to suggest that transfer—near or far—is attained outside of controlled laboratory settings.

What are the conditions under which transfers of different types—both far and near—can occur? In their large-scale qualitative review of WM studies, Morrison and Chein (2011) noted the large heterogeneity among the available studies in observing far transfer effects. In their efforts to isolate factors that might account for the variability in the results, they distinguished between two approaches to WM training: In strategy training trainees are introduced to a specific strategy in the beginning of training (e.g., mental imagery), and provided practice sessions to apply that strategy over sessions. The aim is to teach effective approaches to encoding, maintenance, and retrieval of information from WM. In contrast, core training involves the repetitive administration of demanding WM tasks (e.g., n-back) designed to target domain-general WM mechanisms. Morrison and Chein’s (2011) analysis suggested that core training offers a greater chance of transfer, possibly because it targets domain-general WM mechanisms likely shared across tasks. This point is worth emphasizing: Because structurally different tasks that draw on WM capacity are more likely to share domain-general than domain-specific WM mechanisms, adopting a training approach that targets the former should increase one’s odds for overlapping processes, and therefore success in transfer.

The importance of shared and overlapping WM mechanisms and processes as a factor in transfer has received much attention recently. Harrison et al. (2013) compared the efficacy of training on simple and complex WM span tasks to training on a control task involving visual search. Transfer effects were tested using structurally similar and dissimilar simple and complex WM span tasks, as well as general intellectual function including performance on the RAPM. Their results demonstrated that training on complex WM span tasks led to improved performance on similar tasks (i.e., reading span and rotation span), despite the fact that the target tasks used different material including distractors and to-be-remembered items. However, neither simple nor complex WM span tasks led to gains in fluid intelligence as measured by the RAPM. Not only do Harrison et al.’s (2013) result point to structural similarity between the trained and untrained tasks as a critical variable facilitating transfer, but they go one better by also showing that differences in superficial surface features are not important factors for consideration. Of course, structural similarity implies that the same underlying WM mechanisms are likely engaged by the two tasks.

The same conclusion can be drawn from the study conducted by Dahlin et al. (2008). They demonstrated transfer to a 3-back test of WM after five weeks of training in a specific aspect of WM—updating. Importantly, using fMRI, they were also able to determine that the transfer effect was based on a joint training-related increase in brain activation in the trained and target tasks in the striatum. Importantly, no transfer was observed to a task that did not involve updating, and did not engage the striatum. Dahlin et al.’s (2008) results suggest that to obtain transfer, it is necessary to train specific aspects of WM (e.g., updating) that are functionally shared by
The prospects of cognitive (brain) training

trained and target tasks. Similarly, likelihood of transfer is increased to the extent that training-related changes in brain function occur in the same region that is recruited in relation to the trained process (e.g., updating) in both tasks.

However, determining the structural and functional architecture of WM tasks and thereby their degree of process-specific association is not a trivial task. For example, the n-back task is used widely by researchers as a WM task. It has face validity as a WM task, because it necessitates maintenance and dynamic updating of a rehearsal set. Kane et al. (2007) set out to investigate its construct validity as a WM task by assessing the correlation between performance on the n-back task and a measure of verbal WM span (operation span task). In fact, they found a very weak correlation between the n-back and verbal WM span. Furthermore, n-back and verbal WM span were shown to account for independent variance in fluid intelligence as measured by the RAPM. These findings challenge the view that n-back and verbal WM span measure the same WM processes, despite the fact that this assumption appears to be made implicitly by many researchers. In summary, we need better behavioural tests of the construct validities of our WM tasks in order to isolate the specific capacities that are affected by engaging them.

It has also been argued that to fully account for individual differences in WM capacity in relation to complex span tasks requires measures of (1) primary memory, (2) secondary memory, and (3) attentional control (see Harrison et al., 2013). Primary memory is recruited by operations engaged by passive maintenance of information, whereas secondary memory is recruited by operations involving the retrieval of recently activated information. This tripartite framework could prove useful for process-related categorization of WM tasks in terms of their underlying subcomponents.

Section summary. The current evidence suggests that WM training is likely to boost WM performance and skills. In contrast, it is unlikely to exhibit transfer to other WM tasks unless the trained and untrained tasks share structural (Harrison et al., 2013) and/or processing (Dahlin et al., 2008; see also Beatty et al., 2014) similarities (i.e., WM mechanisms). Therefore, deciding whether WM training should be used as a tool to aid the teaching of thinking when the intended route of effect is improvement in WM capacity will depend on the extent to which the specific WM task and the specific type of thinking targeted share the aforementioned structural and functional similarities. In addition, given that gains in visuospatial WM appear to be better maintained than gains in verbal WM (Melby-Lervåg & Hulme, 2013), modality might be an important factor in adjusting one’s expectations in relation to the long-term benefits of transfer effects.

WM training: transfer to fluid intelligence

In a series of important publications, Susanne Jaeggi and colleagues initially reported that WM training could increase fluid intelligence in adults (Jaeggi et al., 2008), and then extended the effect to a sample consisting of children (Jaeggi et al., 2011). The significance of the malleability of fluid intelligence for applied settings is difficult to overstate: because many aspects of reasoning and problem solving in novel contexts draw on fluid intelligence, performance in those contexts stands to benefit if WM training were proven effective in boosting fluid intelligence. However, overall, the currently available evidence does not support the claim that WM training can reliably increase fluid intelligence (see Melby-Lervåg and Hulme, 2013). Rather, it appears that the likelihood of far transfer from WM training to general cognitive ability including fluid intelligence might be moderated by a number of factors. Here, we focus on two such factors: Task features and individual differences.

In terms of the former, Stephenson and Halpern (2013) conducted an important study in which they had participants complete variations of the n-back task, or a short-term memory task...
(Spatial Matrix Span) as training. In turn, their performance was assessed on four tests of fluid intelligence and four cognitive tests. The four tests of fluid intelligence included (a) RAPM, (b) Cattell’s Culture Fair Test, (c) WASI Matrix Reasoning subtest, which is one of four tests in the WASI—an abbreviated test of the Wechsler Adult Intelligence Scale (WAIS), and (d) BETA-III Matrix Reasoning subtest. The Matrix Reasoning subtest is one of five tests of the BETA-III. The four cognitive tests were (a) Mental Rotation Test, (b) Paper Folding Test, (c) Extended Range Vocabulary, and (d) Lexical Decision Test. Here we focus on the main objective of the study, which was to isolate those features of the training task that facilitate transfer to fluid reasoning. Specifically, participants trained on one of the following four platforms: (a) dual n-back task, (b) visual n-back task, (c) verbal n-back, or (d) short-term memory task training program. The results demonstrated transfer from training to fluid intelligence as measured by RAPM and BETA-III Matrix Reasoning subtest, but critically, only for those participants who trained on the dual n-back task, the visual n-back task, or the short-term memory (STM) task. Therefore, the authors concluded that transfer from WM training to fluid intelligence is contingent on the training task having a visuospatial component.

However, why did training the STM lead to gains in fluid intelligence? The authors offered two explanations. First, it could be that the specific STM task used in their study enhanced a visuospatial mechanism shared by STM and WM. Second, it could be that STM training had an effect because it enhanced the shared short-term storage component that influences fluid intelligence. The main conclusion to be drawn from Stephenson and Halpern’s (2013) study is that the specific features of the training task (and thereby the specific ability that is exercised) is critical for far transfer to fluid intelligence. In other words, Jaeggi et al.’s (2008, 2011) original findings might have been driven specifically by the visuospatial component of their training task (i.e., dual n-back).

The other factor important to transfer might be a set of relevant individual differences variables that might hinder or facilitate it. In a recent study addressing this issue, Jaeggi et al. (2014) compared the impact of two different types of WM training (i.e., single n-back and dual n-back) vs. an active control condition on transfer to factor-analytically derived measures of (fluid) visuospatial and verbal reasoning. In addition, the researchers were also interested in the effects of motivation, need for cognition, pre-existing ability, and implicit theories about cognition on success in training as well as transfer. Here we will focus on three key findings. First, the researchers found that compared to the active control condition, both types of WM training led to gains in visuospatial reasoning but not in verbal reasoning. This replicated earlier results showing transfer of WM training to fluid intelligence. Although it is not clear why a transfer effect was observed for visuospatial but not verbal reasoning, the authors suggested that one reason might be that the former draws more exclusively from fluid intelligence, whereas verbal reasoning draws from both crystallized as well as fluid intelligence, thus possibly diluting the benefits of WM gain on fluid intelligence. Second, they found that participants who reported higher beliefs in the malleability of intelligence showed a greater level of transfer to visuospatial reasoning than did participants who reported believing that intelligence is fixed. Third, motivation was an important factor in training success. Specifically, participants who completed the rigorous training regimen (i.e., one 20–30-minute session per day, five days a week, for a total of 20 sessions) were more likely to report engagement levels that remained relatively constant throughout the four-week period, compared to those who dropped out of training. In fact, there was a significant correlation between self-reported engagement and gain in training. Drop-out is a major issue in WM training studies. For example, in Harrison et al.’s (2013) recent study, 32 of the originally 87 recruited participants (37%) dropped out of the study. Although the authors noted that attrition was unrelated to training condition, the motivational reasons for
dropping out might nevertheless be important in contributing to our understanding of individual differences in gaining from training, and related transfer effects.

Aside from stable (dispositional) individual differences (e.g., implicit theories about intelligence), there are other types of individual differences that must be taken into consideration in evaluating the effect of WM training on transfer to fluid intelligence, such as training performance. For example, Jaeggi et al. (2011) trained elementary and middle school children using a videogame-like WM task. Compared to children assigned to an active control condition, the researchers observed transfer to fluid intelligence only in those children who showed significant improvement on the WM task. In other words, mere assignment to the WM training condition was insufficient for observing transfer. Rather, significant improvement (gain) in training was predictive of transfer. Indeed, the difference between the significantly improved trained group and the control group remained intact even after a three-month hiatus from training, suggesting that the transfer effect was reliable and lasting.

Section summary. The current evidence suggests that WM training does not reliably increase fluid intelligence across the board. However, we are beginning to gain a better understanding of the factors that might facilitate or hinder transfer. First, it has recently been shown that it might be necessary for the training task to have a visuospatial component for transfer to occur (Stephenson & Halpern, 2013). Second, individual differences appear to play a part in how likely one is to show training-related gain on the trained task, as well as exhibit far transfer effects (Jaeggi et al., 2011, 2014). In this sense, researchers and practitioners must develop a better understanding of various task-related as well as motivational, attitudinal, and training-related variables to help predict training success.

WM training and ADHD

Although many different training tools and regimens have been used to study the impact of WM training on outcome measures of interest in children with ADHD, here we will focus on studies that have used Cogmed Working Memory Training (CWMT, www.cogmed.com). CWMT is a computerized task battery designed to target both the storage and manipulation aspects of verbal and nonverbal WM. The subjects typically train for approximately 30–45 minutes per session, and the sessions typically last for four consecutive weeks. Importantly, CWMT is adaptive, meaning that the difficulty level of the task is adjusted as a function of performance. This ensures that the subject is maximally exerted as performance changes across sessions. Prior to conducting their own randomized placebo-controlled study on the effectiveness of CWMT for treating ADHD symptoms in 5–7 year old children, van Dongen-Boomsma et al. (in press) reviewed the results of six earlier studies that had used CWMT to train WM in children with ADHD. Of those studies, five had reported data on cognitive performance, showing improvement on at least one trained WM task. In addition, of the four studies that had reported data on performance on untrained tasks (i.e., far transfer), two studies reported improvement. Finally, and perhaps most importantly for this population, four of the studies also reported data on core behavioural symptoms of ADHD, with three showing training-related improvement. In summary, based on the original six studies it was difficult to infer whether WM training as administered via CWMT is a reliable tool for reducing ADHD symptoms in children, and for observing far transfer effects on outcome measures of interest. However, van Dongen-Boomsma et al. (in press) also noted that drawing firm conclusions was made difficult because of various methodological issues that may have affected the reported outcomes, including variations in time-on-task and parenting/coaching between the treatment and control groups (see also Chacko et al., 2014).
To answer the question of whether CMWT is indeed effective in alleviating ADHD symptoms and improving school performance in young children, van Dongen-Boomsma et al. (in press) and Chacko et al. (2014) conducted two separate studies that addressed many of the key methodological concerns of the previously reported six studies. Van Dongen-Boomsma et al. (in press) used a triple-blind randomized placebo-controlled design in which 51 children (5–7 years old) were assigned to an active CWMT training or a placebo condition, 47 of whom met compliance criteria and were therefore included in the final analysis. Participants in the active CMWT condition trained for 15 minutes per day, five days a week, for a total of 25 sessions. The training package consisted of seven visuospatial WM tasks, wherein difficulty level was adjusted in relation to each child’s performance. Participants in the control condition underwent an exactly similar protocol, except that the WM items to be remembered never exceeded the starting level of two items. All parents were instructed to encourage the children during the course of training, and a certified CWMT coach contacted the parents every week to assess the motivation and performance level of the children. The study was triple blind because neither the children, nor the parents or the coaches were made aware of the treatment condition.

The researchers used three classes of outcome measures. The first class consisted of the core behavioural symptoms of ADHD, rated by the investigators and teachers at baseline and endpoint of training. The second class consisted of neurocognitive measures (e.g., verbal WM, nonverbal reasoning ability, memory for sentences and spatial sequences, etc.). The third class consisted of global clinical functional outcomes, such as ratings by the researchers that the children were “much improved” or “very much improved” at endpoint of training compared to baseline.

The results were largely not supportive of CWMT as an effective intervention strategy for children with ADHD. Specifically, WM training had no effect on the core behavioural symptoms of ADHD or on global clinical functional outcomes. The only effect occurred on verbal WM, although it did not survive correction for multiple comparisons (involving all neurocognitive measures). Van Dongen-Boomsma et al. (in press) concluded that compared to placebo training, there is no robust evidence to suggest that CWMT training confers an advantage for improving cognition or ADHD symptoms in 5–7 year old children.

Chacko et al. (2014) addressed the same major questions as van Dongen-Boomsma et al. (in press) using a very similar design. They randomly assigned 85 slightly older children (7–11 years old) to either an active or placebo CWMT condition. Of those, 80% of the children in the active training condition and 77% of the children in the placebo condition met compliance criteria. As in van Dongen-Boomsma et al.’s (in press) study, the critical difference between the active and placebo conditions was that in the former condition difficulty level was adjusted in relation to each child’s performance. Each child trained for three weeks, and their performance was monitored and encouraged by parents and coaches.

Chacko et al. (2014) focused on four classes of outcome measures. The first class consisted of parent and teacher reports of ADHD symptoms. The second class consisted of a battery of WM measures that tapped its storage (e.g., dot matrix) as well as processing/manipulation (e.g., spatial recall) aspects. The third class consisted of objective assessments of attention, activity level, and impulse control. Specifically, motor activity was recorded using an acceleration-sensitive device (i.e., actigraph) that stored movements per minute. In addition, inattention and impulsivity were assessed using the A-X Continuous Performance Test (Halperin et al., 1991). The fourth class consisted of assessing academic achievements in the areas of word reading, sentence comprehension, spelling, and mathematical computation.

The results demonstrated that there was no difference between the CWMT and placebo groups in parent or teacher reports of ADHD symptoms. Similarly, no treatment effect was
observed on the objective assessments of attention, activity level, and impulse control, or academic achievements in the areas of word reading, sentence comprehension, spelling, and mathematical computation. In contrast, the data from the battery of WM measures demonstrated that the CWMT group improved significantly more than the placebo group on nonverbal and verbal storage in the course of the three weeks.

In conjunction, the results of Chacko et al. (2014) and van Dongen-Boomsma et al. (in press) demonstrate that although WM training as administered by CWMT can bring about improvement in specific aspects of WM (e.g., storage), the evidence does not suggest that this intervention is effective in alleviating ADHD symptoms or improving school performance in children. Reflecting on the results of Chacko et al. (2014), Gathercole (2014) agreed with the conclusion that based on the current body of data WM training should not be recommended as a treatment for ADHD, but stopped short of concluding that WM training has no therapeutic utility. Rather, she argued that the challenge is to design training programs that can generate substantial functional transfer and lead to practical cognitive benefits. She outlined three ways in which this might be achieved. First, hybrid training programs could be developed by combining features of complex span training and n-back training. Such an approach would increase the likelihood that diverse aspects of WM are exercised during training. Second, conventional WM training periods could be augmented with a period of training for transfer designed specifically to apply the acquired skills and abilities to the practical contexts and settings for which it was designed, such as classrooms. Third, to minimize transfer failure, WM training methods could be integrated directly into those activities that children with poor WM capacity show impairments in, such as following instructions, mental arithmetic, and language comprehension. In summary, Gathercole’s (2014) recommendations involve concrete insights for moving the field forward, and refocus the main question away from whether WM training is effective in alleviating ADHD symptoms to the search for conditions that will serve to optimize the therapeutic utility of WM training.

Conclusions

This chapter conducted a selective review of the literature on WM training to determine whether cognitive (brain) training is effective as an aid for teaching thinking, including in children with ADHD. The evidence suggests that whereas WM training can improve performance (and alter related brain function) on the trained and other structurally similar tasks (i.e., near transfer), transfer to structurally dissimilar untrained tasks or general cognitive function (i.e., far transfer) is not reliably observed. Whether far transfer occurs or not could depend on several factors. First, core training appears to offers greater chance of transfer, possibly because it targets domain-general WM mechanisms likely shared across tasks (Morrison & Chein, 2011). This suggests that the specific ways in which one trains is an important factor. In line with the idea that the structural specifics of the training task are critical, Stephenson and Halpern (2013) have recently shown that transfer from WM training to fluid intelligence is contingent on the former having a visuospatial component. Second, whether transfer occurs could be a function of the extent to which the specific abilities (and related brain regions) affected by WM training overlap with those recruited by the untrained tasks (Beatty et al., 2014; Dahlin et al., 2008). In turn, the greater the structural and functional similarities between the trained and untrained tasks, the greater will be the likelihood that those sets of regions will overlap. Third, Jaeggi et al. (2011, 2014) have highlighted the role of individual differences (e.g., motivation, beliefs, etc.) as important factors that moderate the effect of WM training on far transfer. The next steps in this field will likely involve targeted examinations of the specific factors that facilitate transfer—both near and far.
However, despite the fact that reliable far transfer from WM training is not always observed, great opportunities exist for unearthing how training on specific WM tasks (with focus on their underlying processes and mechanisms) can lead to improvements in other WM tasks. Here, the observations of Harrison et al. (2013) are germane:

It is becoming very clear that training on working memory with the goal of trying to increase Gf [fluid intelligence] will likely not succeed (Melby-Lervåg & Hulme, 2013; Shipstead, Redick, & Engle, 2012). More important, this focus might cause one to miss the more realistic goal of training those specific strategies and mechanisms of the working memory system important to other aspects of real-world cognition.

(p. 2418)

In other words, by becoming more precise in our thinking and more modest in our goals, WM training could yet prove effective in improving cognition in practice.

References

The prospects of cognitive (brain) training


