

# Is Working Memory Training Effective?

Zach Shipstead, Thomas S. Redick, and Randall W. Engle  
Georgia Institute of Technology

Working memory (WM) is a cognitive system that strongly relates to a person's ability to reason with novel information and direct attention to goal-relevant information. Due to the central role that WM plays in general cognition, it has become the focus of a rapidly growing training literature that seeks to affect broad cognitive change through prolonged training on WM tasks. Recent work has suggested that the effects of WM training extend to general fluid intelligence, attentional control, and reductions in symptoms of ADHD. We present a theoretically motivated perspective of WM and subsequently review the WM training literature in light of several concerns. These include (a) the tendency for researchers to define change to abilities using single tasks, (b) inconsistent use of valid WM tasks, (c) no-contact control groups, and (d) subjective measurement of change. The literature review highlights several findings that warrant further research but ultimately concludes that there is a need to directly demonstrate that WM capacity increases in response to training. Specifically, we argue that transfer of training to WM must be demonstrated using a wider variety of tasks, thus eliminating the possibility that results can be explained by task specific learning. Additionally, we express concern that many of the most promising results (e.g., increased intelligence) cannot be readily attributed to changes in WM capacity. Thus, a critical goal for future research is to uncover the mechanisms that lead to transfer of training.

*Keywords:* working memory, training, attention, Cogmed, general fluid intelligence

The observed training effects suggest that [working memory] training could be used as a remediating intervention for individuals for whom low [working memory] capacity is a limiting factor for academic performance or in everyday life. (Klingberg, 2010, p. 317)

Fluid intelligence is trainable to a significant and meaningful degree . . . and the effect can be obtained by training on problems that, at least superficially, do not resemble those on the fluid-ability tests. (Sternberg, 2008, p. 6792)

Future research should not investigate *whether* brain training works, but rather, it should continue to determine factors that moderate transfer. (Jaeggi et al., 2011, p. 10085)

Does [working memory] training yield generalized cognitive enhancement? In the case of core training, our answer is a tentative yes. (Morrison & Chein, 2011, p. 57)

The above quotations reflect the growing sentiment that, despite more than 100 years of equivocal results (Carroll, 1993; Jensen, 1998; Thorndike & Woodworth, 1901), psychologists have devised effective methods for training cognitive abilities. This enthusiasm stems from the relatively recent identification of working memory (WM) as a central component of general cognition (cf. Cowan et al., 2005; Engle, Tuholski, Laughlin, & Conway, 1999; Oberauer, Schulze, Wilhelm, & Süß, 2005). Taking advantage of this perspective, many modern training programs are thus designed to specifically target WM (cf. Klingberg, 2010; Sternberg, 2008). In turn, it is assumed that, if a person's WM can be strengthened, a constellation of related abilities will benefit.

This assumption has been reinforced by several studies that have concluded that trained participants are better equipped to reason with novel information (Jaeggi, Buschkuhl, Jonidas, & Perrig, 2008; Jaeggi, Studer-Luethi, et al., 2010; Klingberg et al., 2005; Klingberg, Forssberg, & Westerberg, 2002), have improved attention (Chein & Morrison, 2010; Klingberg et al., 2005, 2002), and, in certain cases, display decreases in ADHD-related symptoms (Beck, Hanson, Puffenberger, Benninger, & Benniger, 2010; Klingberg et al., 2005, 2002; Mezzacappa & Buckner, 2010). Driven by these encouraging results, WM training has rapidly gained prominence within the psychological literature. In recent years, numerous articles have appeared in high-profile journals such as the *Proceedings of the National Academy of Sciences* (Jaeggi et al., 2008; Jaeggi, Buschkuhl, Jonidas, & Shah, 2011), *Science* (E. Dahlin, Stigsdotter Neely, Larsson, Bäckman, & Nyberg, 2008; McNab et al., 2009), *Psychological Science* (Houben, Wiers, & Jansen, 2011; Persson & Reuter-Lorenz, 2008), *Nature Neuroscience* (Olesen, Westerberg, & Klingberg, 2004), *Intelligence* (Colom, Martínez-Molina, Chun Shih, & Santacreu, 2010; Jaeggi, Studer-Luethi, et al., 2010), and *Trends in Cognitive Sci-*

---

Zach Shipstead, Thomas S. Redick, and Randall W. Engle, School of Psychology, Georgia Institute of Technology.

Thomas S. Redick is now at the Division of Science, Indiana University-Purdue University, Columbus.

This work was supported by a grant from the Office of Naval Research (No. N00014-09-1-0129). We thank Kenny Hicks, Tyler Harrison, Kevin Strika, and Forough Azimi for their assistance on various drafts of this article. We also thank Dave Pisoni, Steven Beck, Susanne Jaeggi, Alexandra Morrison, Helena Westerberg, Roberto Colom, Rachael Seidler, Lauren Richmond, Enrico Mezzacappa, Bradley Gibson, Karin Dahlin, and Yvonne Brehmer for providing additional information.

Correspondence concerning this article should be addressed to Zach Shipstead, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332. E-mail: zshipstead@gatech.edu

ences (Klingberg, 2010). In turn, commercial versions of the tasks used in these studies have become readily available (e.g., Cogmed Working Memory Training, 2006; Alloway & Alloway, 2008). These products are promoted as being backed by scientific research and make diverse promises such as improved grades in school (Jungle Memory, 2010), better control of attention and impulses (Cogmed, 2010), and increased IQ (Mindspark, 2011).

While recent reviews have expressed optimism for WM training (Klingberg, 2010; Morrison & Chein, 2011), our concern is that these articles, like much of the literature, have not placed enough emphasis on (a) developing a thorough, empirically based, account of WM; (b) exploring confounds that might account for training effects; and (c) providing detailed analysis of the literature, within the context of the first two concerns. In particular, the tendency to judge the literature solely on the presence or absence of general effects, while neglecting the importance of understanding the ostensible mechanisms of training, is a prevalent oversight that we ourselves have made (i.e., Shipstead, Redick, & Engle, 2010).

Our present intent is thus not to draw conclusions regarding the potential efficacy of WM training. Rather, our goal is to explore the fundamental assumptions that (a) training improves WM and (b) ostensible improvements in cognition can be attributed to WM training. We conclude that these assumptions have yet to be systematically demonstrated. In making this argument, we attempt to highlight specific challenges for future research.

## Working Memory

### Working Memory, Short-Term Memory, and General Cognition

To facilitate understanding of the WM training literature, we first develop a perspective of what WM is and what it is not. In particular, we differentiate WM from the concept of short-term memory (STM).

STM is traditionally thought of as the amount of information a person can simply retain over a brief interval of time. One method for measuring a person's STM is the "simple span" task (cf. Daneman & Carpenter, 1980; Engle & Oransky, 1999). As illustrated in Figure 1a, simple span tasks present a series of verbal (e.g., letters, words, digits) or visuo-spatial (e.g., locations on a grid) items. After the last item, the test-taker is signaled to recreate the list in serial order. Testing begins with short lists (two to three items) that increase in length over the course of several trials. Testing ends when a person is no longer capable of recalling an entire list. Thus, STM can be thought of as temporary storage and is experimentally defined through the longest list of items a person can accurately recall.

In many of the studies reviewed here, WM is explicitly defined as a storage system that is responsible for retaining small amounts of information over brief intervals of time and measured via the above methods (e.g., Klingberg, 2010; McNab et al., 2009; Olesen et al., 2004). However, if storage is the mechanism that relates WM to higher cognition, then occupying a person's STM should severely disrupt reasoning ability.

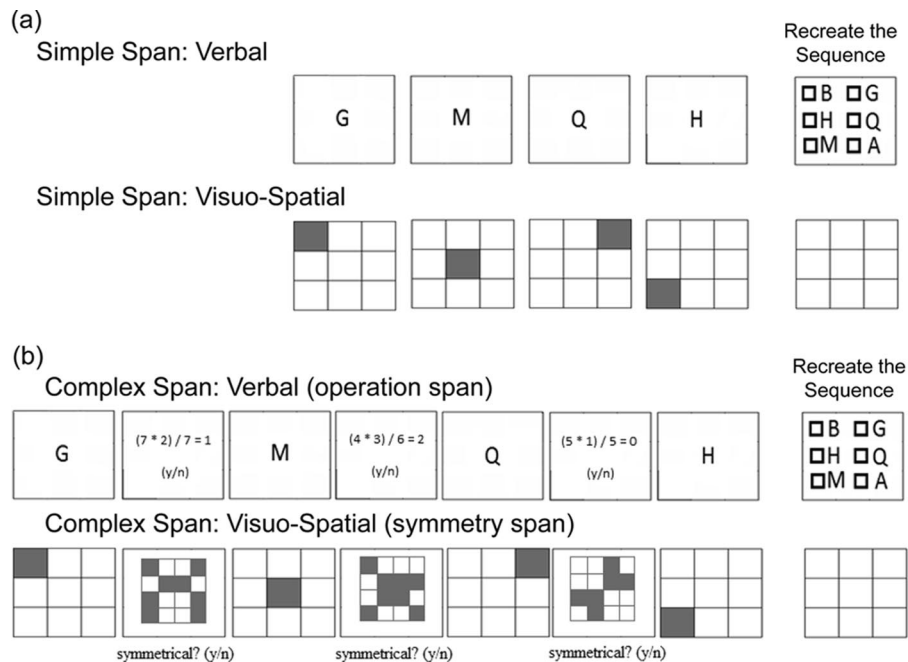


Figure 1. Examples of (a) simple span and (b) complex span tasks. Each box represents information that is presented at a single point in time. Time flows from left to right. In the simple span, test-takers see a series of letters or spatial locations and must recreate the list after the last item is presented. The complex span follows the same procedure, with the exception that a processing task must be completed in between the presentation of each to-be-remembered item.

This hypothesis was directly tested by Baddeley and Hitch (1974), who required participants to perform verbal reasoning tasks while concurrently maintaining two or six digits in STM. This second condition is of particular relevance, since six digits is at the boundary of the average person's STM (Chase & Ericsson, 1982; Cowan, 2001; Miller, 1956; Morey & Cowan, 2005). Thus, attempting to reason while simultaneously remembering this information should have been catastrophic to performance. Instead, participants showed only a slight decline.

Baddeley and Hitch (1974) thus concluded that complex cognition is not dependent upon STM. Rather, it is the other way around. When to-be-remembered information exceeds the capacity of temporary storage, a "general-purpose work space" (now known as the *central executive*) can be engaged to provide support. Thus, STM was demoted to a subcomponent of the larger WM system. This interpretation was further reinforced by inconsistent correlations between individual differences in STM and verbal ability (cf. Crowder, 1982; Daneman & Carpenter, 1980; Perfetti & Lesgold, 1977; Turner & Engle, 1989).

Subsequent tests were developed under the assumption that measurement of individual difference in WM capacity (i.e., the efficacy with which WM functions) should require acts of both storage and processing (Case, Kurland, & Goldberg, 1982; Daneman & Carpenter, 1980; Turner & Engle, 1989). In particular, complex span tasks (Daneman & Carpenter, 1980; Figure 1b) follow a procedure that is similar to simple span tasks, with the exception that test-takers are required to complete a simple processing task (e.g., mathematical operation; symmetry judgment) between the presentation of each item. This requirement of memory in the face of distraction thus increased the role of attention (cf. Engle & Oransky, 1999; Kane, Conway, Hambrick, & Engle, 2007) and retrieval from long-term memory (cf. Healey & Miyake, 2009; Unsworth & Engle, 2007a, 2007b, 2007c; Unsworth & Spillers, 2010).

In contrast to the simple span, the complex span has proven to be a reliable predictor of cognitive ability (Daneman & Merikle, 1996; Unsworth & Engle, 2007b). In particular, WM capacity (as defined by complex span and other WM tasks) is strongly related to a person's ability to reason with novel information (i.e., general fluid intelligence; Gf). Though the exact distinction between WM capacity and Gf is a source of debate (Ackerman, Beier, & Boyle, 2005; Engle, 2002; Heitz et al., 2006; Kyllonen & Christal, 1990; Salthouse & Pink, 2008), it is clear that these constructs are strongly related. Individual differences in WM capacity and Gf show at least a 50% overlap in variation (Kane, Hambrick, & Conway, 2005; Oberauer et al., 2005): A person's ability to reason with novel information can be largely attributed to WM capacity, and vice versa. For this reason, the effect of WM training on Gf is a focus of the literature.

Beyond Gf, there is a well-established relationship between WM capacity and attentional control (Engle, 2002; Kane, Conway, Hambrick, & Engle, 2007; Unsworth & Spillers, 2010; Unsworth, Spillers, & Brewer, 2009). Attentional control refers to the ability to direct attention toward goal-relevant information and away from strong distraction. Relative to people with low WM capacity, individuals with high WM capacity are less likely to have their attention inappropriately drawn into strong distraction, such as hearing their own name (Colflesh & Conway, 2007; A. R. A. Conway, Cowan, & Bunting, 2001), or responding to a peripheral

flash (Hutchison, 2007; Kane, Bleckley, Conway, & Engle, 2001; Unsworth, Schrock, & Engle, 2004). Moreover, individuals with high WM capacity are less apt to mind-wander when focus is needed (Kane, Brown, et al., 2007). Thus, in addition to the potential to improve reasoning ability, WM training may help people become more attentive in their daily activities.

Within the WM training literature, changes to attention are often indexed using the Stroop (1935) task. This task simply requires test-takers to state the hue in which a word has been printed. The difficulty involved in the Stroop task relates to people's tendency to direct their attention toward reading words rather than noticing the ink color in which the words are printed (MacLeod, 1991; MacLeod & MacDonald, 2000). Unsurprisingly, people are faster and more accurate at naming the hue when it is congruent with the word (e.g., the word *BLUE* printed in blue ink), relative to when it is incongruent with the word (e.g., the word *GREEN* printed in blue ink). This difference in time and accuracy is referred to as the Stroop effect.

The size of the Stroop effect is not fixed but partially driven by the ratio of congruent to incongruent trials (Kane & Engle, 2003; Logan & Zbrodoff, 1979; see also Hutchison, 2007). When the task is largely composed of incongruent trials, the Stroop effect shrinks. By the account of Kane and Engle (2003; see also MacLeod & MacDonald, 2000), consistent mismatch between words and hues serves as a continual reminder that reading the words will hurt performance. This constant reinforcement allows attention to be easily directed away from words. However, when the overall proportion of congruent trials is high, the Stroop effect increases (Logan & Zbrodoff, 1979). Under these circumstances, word information generally provides the appropriate response. As a consequence, the need to ignore the word is inconsistently reinforced, and people are apt to lose track of this goal. Thus, it is in these circumstances (i.e., high proportion of congruent trials) that Stroop performance reflects a person's ability to maintain a goal and use it to proactively bias behavior (Kane & Engle, 2003).

It is thus telling that WM capacity is only related to Stroop performance when congruent trials are included in the task (Hutchison, 2007; Kane & Engle, 2003; see also Hutchison, 2011). Individuals with low WM capacity become slow to provide appropriate responses and more prone to overtly reading Stroop words (as indexed by error speed and error type; Kane & Engle, 2003). Thus, the Stroop task reveals a critical aspect of the relationship between WM and attention: WM is not always related to attention (Kane, Poole, Tuholski, & Engle, 2006; Sobel, Gerrie, Poole, & Kane, 2007). Rather it relates to attentional control when prepotent responses must be overcome, particularly in the face of an unsupportive environment.

### Other Relevant Working Memory Tasks

Beyond the complex span task, the *n*-back (Kirchner, 1958; Figure 2a) and running memory span (Pollack, Johnson, & Knaff, 1959; Figure 2b) tasks are also critical to the present discussion. These tasks require test-takers to attend to a serially presented list and remember the most recent three or four items (e.g., letters, spatial locations). However, *n*-back and running span require different types of responses. The *n*-back requires test-takers to make a specific response (e.g., key press) each time the currently presented item matches the item that was presented *n* ago. The

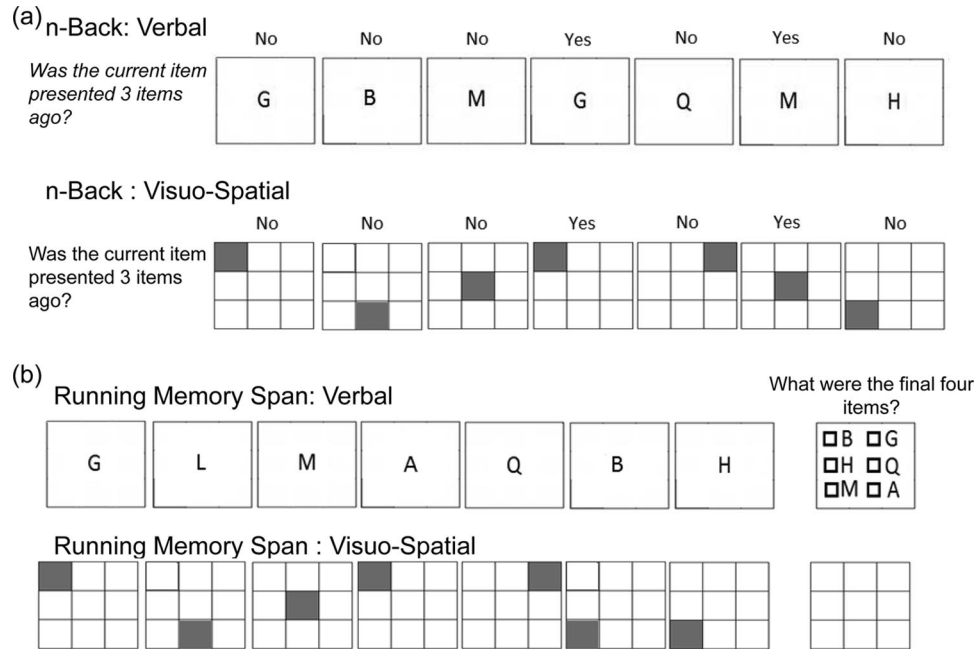


Figure 2. Examples of (a) *n*-back and (b) running memory span. Each box represents information that is presented at a single point in time. Time flows from left to right. In this version of the *n*-back, the test-taker responds “Yes” whenever the currently presented item matches an item that was presented three screens ago. In the running span the test-taker attends to the entire sequence, then is prompted to remember the last *n* items. In this example it is the last four.

challenge of the *n*-back is thus not the length of a list, but the size of *n*. The running span, on the other hand, requires participants to wait until all items have been presented and then recall the last *n* items.

The cognitive mechanisms involved in *n*-back performance are not well understood, and its relationship to the complex span task is unclear (Jaeggi, Buschkuhl, Perrig, & Meier, 2010; Kane, Conway, Miura, & Colflesh, 2007). Nonetheless, the *n*-back does predict individual differences in *Gf* and is generally accepted as a WM task (Gray, Chabris, & Braver, 2003; Jaeggi, Buschkuhl, et al., 2010; Kane, Conway, Miura, & Colflesh, 2007; Schmiedek, Hildebrandt, Lövdén, Lindenberger, & Wilhelm, 2009). The cognitive processes involved in the running span are also subject to debate. Some researchers assume that it taps a person’s ability to update the contents of immediate memory (E. Dahlin, Nyberg, Bäckman, & Stigsdotter Neely, 2008; E. Dahlin, Stigsdotter Neely, et al., 2008; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), others argue that it indexes the amount of information that can be attended in one instant (Bunting, Cowan, & Sauls, 2006; Cowan et al., 2005), and still others argue that it measures the same processes as complex span tasks (Broadway & Engle, 2010). This controversy aside, the foregoing researchers agree that the running span is a valid WM task and, despite the lack of an interpolated processing task, highly related to the complex span (Broadway & Engle, 2010; Cowan et al., 2005; Miyake et al., 2000).

Finally, for younger children, simple span tasks can measure WM capacity, provided the children are required to recall items in reverse order (Alloway, Gathercole, & Pickering, 2006; Engel de Abreu, Conway, & Gathercole, 2010; Gathercole & Pickering,

2000; St. Clair-Thompson, 2010). However, within the WM training literature, researchers sometimes report an average of forward and backward simple span scores, thus diluting the role of WM for children. For this reason, we treat cases in which forward and backward recall have been combined as instances of STM. It is also worth noting that, for adults, simple span performance reflects STM, regardless of whether forward or backward recall is required (Engle et al., 1999; St. Clair-Thompson, 2010).

### Working Memory Beyond the Laboratory

WM capacity has found natural extension to real-world behavior. A cornerstone of early research with young adults was the discovery that WM tasks reliably predict reading comprehension (Daneman & Carpenter, 1980; Turner & Engle, 1989) and performance on college entrance exams (e.g., ACT, SAT; Cowan et al., 2005; Turner & Engle, 1989). Since that time, individual differences in WM capacity have been linked to diverse skills such as learning computer languages (Shute, 1991), sight-reading music (Meinz & Hambrick, 2010), multitasking (Bühner, König, Prick, & Krumm, 2006; Hambrick, Oswald, Darowski, Rench, & Brou, 2010), and regulating emotion (Kleider, Parrott, & King, 2010; Schmeichel, Volokhov, & Demaree, 2008).

WM can be distinguished from STM in children as young as 4 years (Alloway, Gathercole, & Pickering, 2006). By kindergarten, WM capacity is predictive of *Gf* (Engel de Abreu et al., 2010), and individual differences in WM capacity predict children’s verbal and mathematical aptitude (Cowan et al., 2005; Gathercole & Pickering, 2000; Swanson & Beebe-Frankenberger, 2004). For

example, WM capacity has been linked to the rate at which children develop syntactic knowledge and reading ability (Engel de Abreu, Gathercole, & Martin, 2011).

It is critical for the present discussion to note that children with low WM capacity are prone to learning disabilities (Swanson, 2003) and have difficulty carrying out complex instruction (Engle, Carullo, & Collins, 1991; Gathercole, Durling, Evans, Jeffcock, & Stone, 2008; see also Gathercole & Alloway, 2004). A specific area of interest to training researchers is attention-deficit/hyperactivity disorder (ADHD). Although the distinction between ADHD as it relates to deficits of attention and presence of hyperactivity is controversial (Diamond, 2005; Nigg, 2010), Diamond (2005) argued that low WM capacity is symptomatic of attention deficits. From her perspective, ADHD-diagnosed children who specifically show symptoms of attention deficits (absent hyperactivity) are experiencing WM-related deficits of selective attention (i.e., attentional control). Hyperactivity, on the other hand, is assumed to stem from difficulty with response inhibition.

This perspective (i.e., Diamond, 2005) suggests that effective WM training should specifically alleviate deficits of attention, but not tendencies toward hyperactivity. Westerberg, Hirvikoski, Forssberg, and Klingberg (2004), however, have demonstrated that, regardless of presenting-symptoms, ADHD is generally associated with visuo-spatial memory deficits (see Figure 1a), as well as difficulty in consistently allocating attention to a task. Thus, an intervention that strengthens the core functions of WM (which influence STM and attention) may be beneficial to all individuals with ADHD.

## Training Working Memory

### The Adaptive Working Memory Training Paradigm

The identification of WM as a central component of cognition has introduced the possibility that focused, theoretically motivated techniques for training broad abilities can be developed. Klingberg (2010) hypothesized that three factors are critical to a successful cognitive training program (see also Jaeggi, Studer-Luethi, et al., 2010; Morrison & Chein, 2011).

First, training should not teach specific strategies for simply remembering more information (e.g., rehearsal techniques or mnemonic devices). Strategies might improve a person's score on a WM test (McNamara & Scott, 2001; Turley-Ames & Whitfield, 2003); however, this is not the same as changing the underlying ability. In particular, memory strategies tend to be context specific (Chase & Ericsson, 1982; Maguire, Valentine, Wilding, & Kapur, 2003). A rehearsal technique that helps a person remember a series of digits will be useless when applied to nonverbal materials (e.g., snowflakes; Maguire et al., 2003). Moreover, people with cognitive deficits tend to have difficulty recognizing situations in which a strategy might apply (Butterfield, Wambold, & Belmont, 1973). Perhaps most important, when a group of people are all taught the same strategy, scores on WM tests actually become more predictive of cognitive ability (Turley-Ames & Whitfield, 2003): Rather than accounting for individual differences in WM capacity, the varied strategies people use when taking WM tests obstruct accurate measurement by introducing unsystematic variance.

Second, Klingberg (2010) argued that the training program should be specifically focused on WM tasks. The inclusion of

other types of training will be time consuming and thus dilute the efficacy of the intervention. This further implies that WM tasks do not simply measure WM capacity but also stimulate neural plasticity (Klingberg, 2010; Lövdén, Bäckman, Lindenberger, Schaefer, & Schmiedek, 2010; McNab et al., 2009). If enough time is spent on these tasks, WM capacity will increase.

Third, training schedules should be rigorous (roughly 20 sessions, each lasting 30–60 min) and training programs should adapt to user performance. If a person is meeting specific performance criteria, task difficulty should increase. When these criteria are not met, task difficulty should decrease. Thus, trainees are constantly engaged in the task at a level that is neither boring nor overtaxing (Lövdén et al., 2010; e.g., Csikszentmihalyi, 1990). Computerized training is therefore preferred to one-on-one sessions, as it can be conducted in various locations and can automatically adapt to performance.

These factors are built into “adaptive” WM training. Adaptive training tasks are modified versions of standard memory tasks such as the simple and complex span tasks displayed in Figure 1. The difficulty of these adaptive tasks is tied to list length. If a trainee is performing well, the list length increases by one item. If a trainee is struggling, the list length decreases by one item. For an adaptive *n*-back, the size of *n* adjusts, rather than the length of the list. For the adaptive running span, list length increases, ostensibly forcing participants to update the contents of immediate memory more times per trial.

### Transfer of Training

Over the course of 4–5 weeks of training, people typically advance through several levels. However, improved performance on the training task does not signal an increase in WM capacity. For instance, Chase and Ericsson (1982) reported a participant (S.F.) who, after several months of training on an adaptive span task, was able to recall sequences of more than 80 digits. However, when the digits were presented at a faster rate, his scores returned to normal levels. The reason for this decline was that S.F. had developed a strategy of mapping short sequences of numbers onto preexisting knowledge (i.e., cross-country running times, historical dates). When the testing conditions were changed, his strategy could not be employed.

It must, therefore, be demonstrated that the effects of training transfer to untrained tasks (Barnett & Ceci, 2002; Klingberg, 2010). This is typically accomplished within a pretest–posttest design (Campbell & Stanley, 1963; Shadish, Cook, & Campbell, 2002). The pretest consists of a battery of tasks, each of which is designed to measure an ability of interest. This is followed by assignment of participants to either perform several weeks of WM training or serve in a control group. Within a few days of finishing the training regimen, participants complete a posttest in which alternate versions of these tasks are administered.

Posttest improvement on these tasks (relative to both pretest and the control group) thus provides evidence that training has transferred. Improved performance on tasks that are intended to measure WM capacity are termed *near transfer*. Posttest improvement on tasks that are intended to measure related abilities are termed *far transfer*. Near transfer thus provides evidence that WM capacity has increased, as well as a mechanism through which far transfer results may be explained. In the absence of increased WM

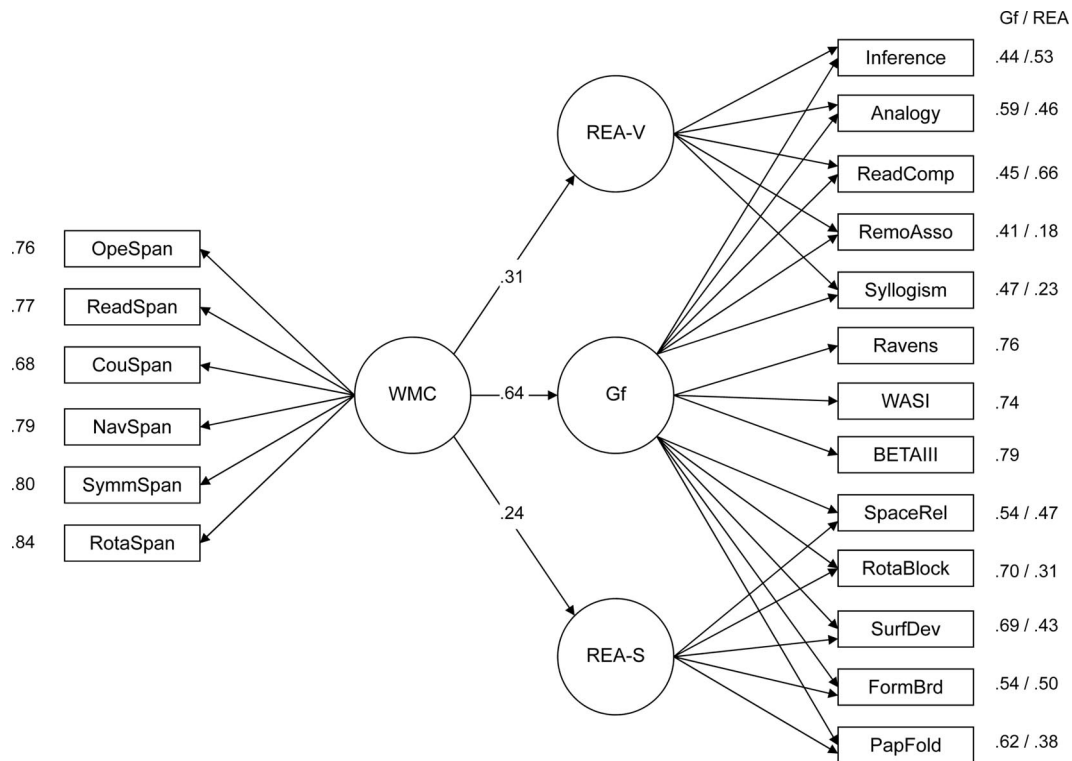
capacity, it is theoretically unclear why WM training should lead to improvements on far transfer tasks.

Pretest–posttest designs are powerful methods for removing confounds (Campbell & Stanley, 1963; Shadish, Cook, & Campbell, 2002). However, they do not control all confounds, and demonstrating valid transfer is more difficult than the above discussion implies. In our own reading of the literature, we have identified four concerns that deserve elaboration.

**Inadequate measurement of abilities.** The goal of cognitive training is to change an underlying ability that is thought to be driving performance of a class of tasks (McArdle & Prindle, 2008). If an intervention has increased a person’s WM capacity, then improvements should not depend on the type of WM task used to measure transfer. However, in most studies, abilities are measured via single tasks. The inherent problem is that scores on any single test are driven both by the ability of interest and other systematic

and random influences (cf. Kim & Mueller, 1978; Loehlin, 2004). Thus, when transfer of training is measured via single tests (as is generally the case), posttest improvements represent the possibility that an underlying ability has changed but do not provide definitive evidence (McArdle & Prindle, 2008; Moody, 2009; Schmiedek, Lövdén, & Lindenberger, 2010; Sternberg, 2008).

The distinction between individual tasks and underlying abilities is illustrated in Figure 3, which presents a structural equation model originally reported by Kane et al. (2004). Boxes in this figure represent individual tasks, while circles represent sources of variation that are common to these tasks. That is, the circles represent factors (what we have termed *abilities*). The left-pointing arrows indicate that the factor labeled WMC (i.e., working memory capacity) directly contributes to the performance of six separate WM tasks. The factor loadings (numbers to the left of each task) indicate that the portion of each task that is explained by WM



*Figure 3.* Structural equation model originally reported by Kane et al. (2004). Numbers to the left of each working memory capacity task represent its loading on the latent WM capacity factor (WMC). Numbers to the immediate right of each reasoning task represent its loading on the latent Gf factor. The next number represents a task’s loading on either the latent REA-V or REA-S factor. Note on tasks: OpeSpan = operation span; ReadSpan = reading span; CouSpan = counting span; NavSpan = navigation span; SymmSpan = symmetry span; RotaSpan = rotation span; Inference = ETS Inferences; Analogy = AFOQT Analogies; ReadComp = AFOQT Reading Comprehension; RemoAsso = Remote Associates; Syllogism = ETS Nonsense Syllogisms; Ravens = Ravens Advanced Progressive Matrices; WASI = Wechsler Abbreviated Scale of Intelligence, Matrix Test; BETAI = Beta III, Matrix Test; SpaceRel = DAT Space Relations; RotaBlock = AFOQT Rotated Blocks; SurfDev = ETS Surface Development; FormBrd = ETS Form Board; PapFold = ETS Paper Folding. Note on latent factors: WMC = working memory capacity; Gf = general fluid intelligence; REA-V = verbal reasoning; REA-S = spatial reasoning. Reprinted from “The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span Reasoning” by M. J. Kane et al., 2004, *Journal of Experimental Psychology: General*, 133, p. 205. Copyright 2004 by the American Psychological Association.

capacity ranges from 46% to 71% (obtained by squaring the factor loadings). While some tasks provide stronger reflections of WM capacity than others, no single task reflects WM capacity alone.

However, WM capacity is not the only feature that is common to the WM tasks in Figure 3. All six are complex span tasks. Thus, while they measure WM capacity, they also require dual-task coordination. Complex span tasks therefore also measure a person's ability to successfully switch between two tasks. This latter influence is not present in many other WM tasks (cf. A. R. A. Conway, Getz, Macnamara, & Engel de Abreu, 2010), such as the *n*-back or running span. Moreover, Kane et al. (2004) reported that, above and beyond these common influences (i.e., WM capacity and dual-task performance), several of the WM tasks in Figure 3 were also individually correlated through similarity of memory items or similarity of processing tasks. Thus, when examining near transfer results, it is important to recognize that improved performance on a WM task does not require an increase in WM capacity. Instead, near transfer may be driven by simply practicing certain types of tasks (e.g., complex span tasks), or certain aspects of tasks (e.g., memory for letters, memory for spatial locations). For instance, some training programs (e.g., Cogmed) include backward span tasks. Thus, when trainees improve their posttest performance on backward span tasks, it is unclear whether this represents increased WM capacity, or the effect of practicing a specific type of information transformation.

Likewise, far transfer tasks are not perfect measures of ability. In many training studies, Raven's Progressive Matrices (Ravens; Raven, 1990, 1995, 1998) serves as the sole indicator of Gf. This "matrix reasoning" task presents test takers with a series of abstract pictures that are arranged in a grid. One piece of the grid is missing, and the test taker must choose an option (from among several) that completes the sequence. Jensen (1998) estimates that 64% of the variance in Ravens performance can be explained by Gf. Similarly, Figure 3 indicates that in the study of Kane et al. (2004), 58% of the Ravens variance was explained by Gf.

It is clear that Ravens is strongly related to Gf. However, 30%–40% of the variance in Ravens is attributable to other influences. Thus, when Ravens (or any other task) serves as the sole indicator of far transfer, performance improvements can be explained without assuming that a general ability has changed. Instead, it can be parsimoniously concluded that training has influenced something that is specific to performing Ravens, but not necessarily applicable to other reasoning contexts (Carroll, 1993; Jensen, 1998; Moody, 2009; Schmiedek et al., 2010; te Nijenhuis, van Vianen, & van der Flier, 2007).

Exactly what that something is may not be intuitively clear, since the processes that tasks measure are not always apparent. For example, Jaeggi et al. (2008) demonstrated far transfer to Gf using the *Bochumer Matrizen-Test* (BOMAT; Hossiep, Turck, & Hassella, 1999). This task is similar to Ravens, with the exception that BOMAT presents information in a 15-item rather than nine-item grid. Moody (2009) noted that standard administration of BOMAT allows for 45 min, while Jaeggi et al. (2008) only allowed 10 min. Moody argues that a strong memory component (i.e., 15 items to remember) coupled with strict time limits (which prevented participants from reaching the more challenging problems) may have increased reliance on temporary memory.

In effect, participants who had received memory training may have been at an advantage. This is ostensibly because they were

better prepared to hold information in a readily available state, as opposed to repeatedly scanning the matrix visually. This would not represent transfer to Gf (i.e., novel reasoning, independent of context), since these effects would not be present if more time were given or if a less memory-dependent reasoning task were used. Preemption of criticisms such as Moody's (2009) is, however, readily accomplished through demonstration of transfer to several measures of an ability.

Unfortunately, the practice of equating posttest improvement on one task with change to cognitive abilities is prevalent within the WM training literature (cf. Jaeggi et al., 2008; Klingberg, 2010). This is partially driven by the time and monetary costs associated with conducting multisession, multiweek studies. Regardless, training studies can greatly improve the persuasiveness of their results by measuring transfer via several tasks that differ in peripheral aspects but converge on an ability of interest (e.g., a verbal, Gf, and spatial task from Figure 3). If a training effect is robust, it should be apparent in all tasks.

#### **Conflation of working memory with short-term memory.**

In many studies, near transfer to WM is measured via simple span tasks (e.g., Figure 1a). Thus, reported increases in WM capacity are often obtained using tasks that are traditionally associated with STM (cf. Daneman & Carpenter, 1980; Engle & Oransky, 1999).

These tasks are not always poor measures of WM capacity. Simple span performance reflects WM capacity, provided (a) testing continues regardless of whether the test-taker recalls the entire list, and (b) partial credit is then assigned for remembering items in their original serial position (Unsworth & Engle, 2006, 2007b). When this methodology is applied, simple span tasks will reflect people's ability to recall information from outside of STM, rather than measuring immediate retention (Unsworth & Engle, 2007b). However, this method is generally not employed in the WM training literature; rather, testing ends when entire lists cannot be properly recalled.

The ramifications of scoring simple span tasks in an all-or-none manner are clarified in Figure 4, which presents data originally reported by Engle et al. (1999). In this study, WM capacity was defined by three complex span tasks (displayed to the left), while STM was defined by three simple span tasks (two that required forward recall and one that required backward recall of list items). After these factors were formed, several other measures of various cognitive functions were allowed to freely load on either. This is represented via the lines extending from the factors to the tasks on the right. Significant paths (solid lines) exist between WM capacity and measures of verbal reasoning (ABCD; Kyllonen & Christal, 1990), recall from long-term memory (immediate free recall from secondary memory; Tulving & Colotla, 1970), and memory updating (keeping track; Yntema, 1963). Thus, the influence of WM capacity extends to several measures of cognitive function. It is reasonable to expect that increasing WM capacity would improve a person's performance on these tasks.

The direct influence of STM, on the other hand, was limited to performance on a short-term retention task (continuous opposites; Kyllonen & Christal, 1990). In essence, the influence of STM on the verbal reasoning, recall, and updating tasks was mediated by WM capacity. Thus, improved performance on all-or-none simple span tasks does not explain far transfer. These tasks are better measures of STM, which does not have broad influence on cognitive abilities.

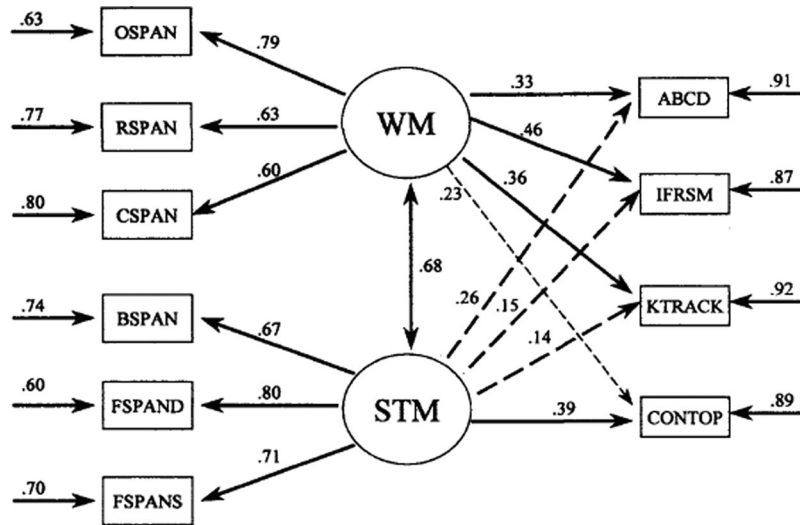


Figure 4. Structural equation model originally reported by Engle et al. (1999). Paths that are significant ( $p = .05$ ) are indicated by solid lines. Broken lines represent nonsignificant paths. Task loading on latent factors is indicated by number above each path. Arrows not associated with a latent factor represent variation in task performance that is not explained by the latent factor. Note on tasks: OSPAN = operation span; RSPAN = reading span; CSPAN = counting span; BSPAN = backward span; FSPAND = forward span/dissimilar; FSPANS = forward span/similar; ABCD = ABCD task; IFRSM = immediate free recall secondary memory; KTRACK = keeping track; CONTOP = continuous opposites. Note on latent factors: WM = working memory; STM = short term memory. Reprinted from "Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach," by R. W. Engle et al., 1999, *Journal of Experimental Psychology: General*, 128, p. 332. Copyright 1999 by the American Psychological Association.

Klingberg (2009) argued that the above discussion is only applicable to verbal STM. As evidence for this position, he points to strong correlations that are often found between visuo-spatial simple span tasks (see Figure 1) and Gf tasks (e.g., Kane et al., 2004; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Shah & Miyake, 1996). This is assumed to reflect a close relationship between Gf and the visuo-spatial domain (Bergman Nutley et al., 2011). However, this correlation is not straightforward.

Returning to the model of Kane et al. (2004; Figure 3), these researchers also found that a factor composed of six visuo-spatial span tasks was strongly correlated with Gf (.54). However, examination of Figure 3 reveals a bias in the way Gf was defined. Of the reasoning tasks employed, five were specifically designed to measure spatial reasoning (i.e., tasks loading on REA-S), while the three tasks that strictly loaded on Gf were all spatially arranged matrix reasoning tasks. Thus, of the 13 tasks contributing to the Gf factor, eight contained strong visuo-spatial components. Kane et al. subsequently redefined the Gf factor as three verbal tasks (i.e., tasks loading on REA-V, verbal reasoning; see Inference, Analogy, and ReadComp in Figure 3) and three visuo-spatial tasks (Figure 3: SpaceRel, RotaBlock, and PaperFold). With this balanced factor, the correlation between visuo-spatial STM and Gf dropped slightly to .47. Next, the Gf factor was defined such that it was verbally biased (Figure 3: Inference, Analogy, ReadComp, and RemoAsso). Under these circumstances the correlation between visuo-spatial STM and Gf dropped to .29. The correlation between WM capacity and Gf, on the other hand, remained largely unchanged across these models.

Thus, visuo-spatial STM tasks introduce a confound. While these tasks measure Gf (to an extent), they also reflect the visuo-

spatial components that are inherent in many Gf tasks (e.g., spatially arranged matrices). As with verbal simple span tasks, visuo-spatial STM tasks should not be expected to reflect general abilities. Rather, their correlation with Gf tasks is inflated by the visuo-spatial format in which Gf tasks are typically presented.

**What type of control group is being used?** Several WM training studies have been conducted without a control group. This should always be a concern, since control groups are critical to eliminating test-retest effects (Cane & Heim, 1950; te Nijenhuis et al., 2007) as well as other experimental confounds (Campbell & Stanely, 1963; Shadish, Cook, & Campbell, 2002). However, a more subtle concern is the prevalent use of "no-contact" control groups, who participate in pre- and posttest sessions but are not otherwise engaged in the experiment. While these groups control for test-retest effects, they simultaneously introduce a new set of concerns.

We have elsewhere (Shipstead et al., 2010) used the term "Hawthorne effect" (French, 1953) to refer to the assumption that people's behavior will be affected by their level of involvement within an experiment. Though the Hawthorne effect traditionally refers to a tendency for people to change their behavior when they know they are being watched (cf. Mayo, 1933), we actually use it to refer to any number of psychological phenomena that may be introduced when training and control groups are treated differently. For instance, control groups may realize that they are not expected to show pretest-posttest improvement (i.e., demand characteristics; Orne, 1962, 1972), while trained participants may come to see themselves as personally invested in improving (i.e., cognitive dissonance; Festinger, 1957). Moreover, the expectations of test administrators can influence outcomes in diverse areas such as



individual differences in classroom performance (i.e., Pygmalion effects; Rosenthal, 1994; Rosenthal & Jacobson, 1968) and the presence of parapsychological effects (Wiseman & Schlitz, 1997). In other words, Hawthorne phenomena introduce a general effect of the influences contained within an experiment (Klingberg, 2010; McCarney et al., 2007; Oken et al., 2008; Shipstead et al., 2010).

One method for controlling Hawthorne/placebo effects has been to administer nonadaptive versions of the training tasks. These tasks never present participants with lists of more than three items, yet keep the control group actively engaged in the experiment. While this is among the best types of control groups that are currently employed, there are reasons for concern.

Placebo effects are controlled by convincing all participants that they are receiving treatment (Rosnow & Rosenthal, 1997). In this regard, the experiences of an adaptive-training group and a non-adaptive control group are not equal. During the training period, the adaptive-training group receives constant feedback via a task that is responsive to their performance. The nonadaptive control group, on the other hand, repeats the same procedure throughout training. Thus, adaptive and nonadaptive groups are treated differently both in terms of rigor of practice (the intended difference) and in terms of being presented with tangible evidence that their abilities are changing (a presumably unintentional difference). Therefore, when training is conducted using adaptive WM tasks, true control requires that both groups be given adaptive tasks: One that involves WM (training group) and one that does not (control group). Thus, any effect of training can be directly attributed to training WM rather than to peripheral experiences within the laboratory.<sup>1</sup>

**Are subjective measures used?** A final concern regards the use of subjective reports as measures of transfer (e.g., Beck et al., 2010; Klingberg et al., 2005; Mezzacappa & Buckner, 2010; Westerberg, Brehmer, D'Hondt, Söderman, & Bäckman, 2008; Westerberg et al., 2007). The laudable goal of this practice is to extend training-related effects beyond the laboratory and into everyday life (Klingberg, 2010). However, an undesirable consequence of subjective reports is the potential for posttest change to reflect expectations that were created by the act of receiving treatment rather than by actual changes that were brought about by the treatment (Aiken & West, 1990; M. Conway & Ross, 1984; DeLoache et al., 2010; Greenwald, Spangenberg, Pratkanis, & Eskenazi, 1991; Orne & Scheibe, 1964).

Greenwald et al. (1991) provided a useful demonstration of the problems associated with subjective reports. Participants in this study received commercially produced audiotapes that contained subliminal messages intended to improve either self-esteem or memory. Unknown to the participants, half of the tapes that were designed to improve memory were relabeled “self-esteem” and vice versa. At a 5-week posttest, participants’ scores on several standard measures of self-esteem and memory were improved, but this change was independent of the message and the label on the audiotape (i.e., participants showed across the board improvement). However, in response to simple questions regarding perceived effects, roughly 50% of participants reported experiencing improvements that were consistent with the label on the audiotape, while only 15% reported improvements in the opposite domain. The self-report measures were neither related to actual improvements in transfer task performance nor related to the content of the

intervention. Instead, they were attributable to expectation of outcome.

## Review of Working Memory Training Studies

With the above concerns in mind, we review the WM training literature across several populations. We focus on training techniques that approximate the methodology outlined by Klingberg (2010; see above) and other researchers (i.e., Jaeggi, Studer-Luethi, et al., 2010; Morrison & Chein, 2011). Several other types of intervention may affect WM capacity. These include mindfulness or meditation training (e.g., Fabbro, Muzur, Bellen, Calacione, & Bava, 1999; van Vugt & Jha, 2011; Zeidan, Johnson, Diamond, David, & Goolkasian, 2010; see also Tang & Posner, 2009), neurofeedback (Cannon et al., 2006; Vernon et al., 2003), physical exercise (Lachman, Neupert, Bertrand, & Jette, 2006), long-term training on musical instruments (George & Coch, 2011; Jones, 2007), and learning various skills (Lee, Lu, & Ko, 2007). We justify our specific focus on computerized, (typically) adaptive, non-strategy-based training by noting that these studies (a) are designed around clearly defined training protocols, thus allowing for comparison across studies, and more importantly (b) are the basis of strong claims regarding the malleability of memory and intelligence (e.g., Mindsparke, 2011; Morrison & Chein, 2011; Sternberg, 2008).

Our discussion of transfer effects primarily focuses on those that are both prevalent within the literature and applicable to the fundamental aspects of WM. Thus, while our discussion focuses on WM, Gf, and attention, we acknowledge that individual articles are not always as restricted in scope.<sup>2</sup> Due to the specificity of population, we do not directly discuss articles that focus on substance abuse (Bickel, Landes, Hill, & Baxter, 2011; Houben et al., 2011), stroke recovery (Westerberg et al., 2007), schizophrenia (Bell, Bryson, & Wexler, 2003; Wexler, Anderson, Fulbright, & Gore, 2000), or multiple sclerosis (Vogt et al., 2009). Similarly, due to differences in training technique, as well as theoretical approach, we restrict our review to studies involving WM training (both adaptive and nonadaptive). Other cognitive training methods such as dual-task performance (Bherer et al., 2005), task-switching (Korbach & Kray, 2009), and practice on attention tasks (Rueda, Rothbard, McCandliss, Saccamanno, & Posner, 2005; Sohlberg, McLaughlin, Pavese, Heidrich, & Posner, 2000) are not directly discussed. The present review should, nonetheless, be applicable to these articles.

We acknowledge that some readers would prefer a review that focuses on the effect size of training results. However, we contend that, as of now, such analyses are potentially misleading in two ways. First, in the present training literature WM capacity is often narrowly defined via transfer tasks that are highly similar to the

<sup>1</sup> For an excellent example of control, see Persson and Reuter-Lorenz (2008). These researchers hypothesized that memory-interference was the source of training effects and arranged their stimuli to either maximize or minimize retrieval competition. This article, however, was recently retracted (Persson & Reuter-Lorenz, 2011), due to the discovery of experimenter error. As such, it is not included in our review. The general methodology is, however, noteworthy.

<sup>2</sup> Here we focus on integration across a circumscribed set of behavioral results across studies. See Shipstead et al. (2010) for a discussion of study-specific outcome measures, as well as discussion of the physiological results reported by Olesen et al. (2004) and McNab et al. (2009).

method of training. Thus, near transfer effects are potentially inflated by task-specific practice and therefore not necessarily interpretable as change to WM capacity. Second, the tasks that some researchers have used to measure change to certain abilities are not always appropriate. For instance, we argue that, due to the exclusion of congruent trials, performance on Stroop tasks used by some researchers (e.g., Klingberg et al., 2005, 2002; Olesen et al., 2004; Westerberg et al., 2008) may not be related to a person's WM. In these cases, the presence or absence of an effect would be equally misleading. If WM capacity is not related to performance on Stroop tasks that exclude congruent trials (Hutchison, 2007; Kane & Engle, 2003), then increased WM capacity cannot readily explain improved performance on such tasks. Thus, we focus our review on integrating studies in terms of the breadth of tasks included in training batteries, and the relationship of near and far transfer results. Through this method of analysis we aim to advance the theory and methodology employed in WM training studies.

A list of relevant Gf and attention tasks is included on Table 1. Additional information regarding the reliability of these and other tasks can be found in the Appendix.

The majority of studies have been conducted using Cogmed Working Memory Training (2006) software. Cogmed training involves several verbal and visuo-spatial simple span tasks that have been embedded within simple video games. Some games involve static displays (e.g., a grid such as Figure 1a), while other games require participants to track movement (e.g., floating asteroids, rotating grids). Cogmed is an adaptive task, in that trial-by-trial performance determines how much information a trainee is required to remember. Both forward and backward recall are practiced. This regimen was developed by Klingberg and colleagues (Klingberg et al., 2005, 2002) and is now commercially available through private practices (Cogmed, 2011). Many studies that use this software were conducted by researchers who are not affiliated with Cogmed. Thus, we do not view the high proportion of studies that use Cogmed software as a weakness. Instead, it allows us to view a training technique that has been applied across several contexts. Additionally, a minority of the reviewed studies are either unpublished (Seidler et al., 2010; Westerberg et al., 2008<sup>3</sup>) or published as book chapters (Shavelson, Yuan, & Alonzo, 2008). We acknowledge that they have not been subjected to peer review. However, we justify the inclusion of these studies as a necessary component of avoiding the file-drawer phenomenon, in which positive results are published while negative results go unnoticed.

## Working Memory Training and Children

Studies that focus on children from populations associated with WM capacity deficits are examined separately from studies that are focused on typically developing children (see Table 2). However, it is clear that independent of the population studied, WM training improves performance on STM tasks. To reiterate, in studies involving children, we define STM as either forward simple span recall or an average of forward and backward performance. Of the 11 studies that fit this rubric (see Table 2), eight report unequivocal transfer, while three (Bergman Nutley et al., 2011; Holmes, Gathercole, & Dunning, 2009; Van der Molen, Van Luit, Van der Molen, Klugkist, & Jongmans, 2010) report mixed transfer (either visuo-spatial or verbal STM).

**Preview of results.** Several studies include valid measures of WM capacity (see Table 2); however, as we discuss, there is concern regarding consistent use of appropriate control groups. In terms of far transfer, some studies that use Cogmed training find evidence of transfer to Gf and attention. However, this transfer has been demonstrated using a limited number of tasks, some of which have questionable relationships to WM. A limited number of studies use *n*-back or running span training and do report associations between progress during training and increased Gf. However, near transfer tasks are not included. Finally, evidence of improvements to ADHD-related symptoms is sparse.

**Working memory training and children with ADHD/low WM capacity.** All studies that were focused on children with ADHD/low WM capacity used simple/complex span tasks at their primary method of training. Cogmed software is most prevalent (see Table 2). Exceptions include Alloway (in press) and Van der Molen et al. (2010), who trained children on Jungle Memory (Alloway & Alloway, 2008) and OddYellow (Van der Molen et al., 2010). Jungle Memory features three adaptive complex span tasks that require memory of visuo-spatial or verbal information (locations or numbers) in the face of interpolated processing tasks (word completion, mental rotation, or mathematics). OddYellow presents trainees with three shapes that differ in terms of the features they possess: Two are identically shaped and two are black. The trainee is required to rapidly (< 5 s) indicate which item has the unique shape and then indicate (< 2 s) which item is yellow. After one to seven trials, the participant must recall the spatial locations of all previously presented yellow items.

**Transfer to WM capacity.** Seven studies that focused on children who were assumed to have WM capacity deficits explicitly tested WM capacity (see Table 2). While Van der Molen et al. (2010) were the only researchers who failed to find any transfer effects, the overall results must be interpreted cautiously. Holmes et al. (2010), Kronenberger, Pisoni, Henning, Colson, and Hazzard (2011), and Mezzacappa and Buckner (2010) did not include a control group; thus, their transfer findings are confounded by repeated testing. K. I. E. Dahlin's (2011) results were relative to the control group of Klingberg et al. (2005).<sup>4</sup> Thus, due to a difference of group characteristics (special education vs. ADHD) and preexisting knowledge of the findings of Klingberg et al. (2005), this result may also be interpreted as a test-retest effect. Finally, Alloway (in press) did not report whether the participants in the training group showed significant improvement on their WM measure. Rather, she only reported that the training group showed larger posttest improvement than did the control group. The concern is that the control group performed numerically worse at posttest than pretest, and this may have been the source of reported transfer.

<sup>3</sup> We focused on behavioral data relative to a control group. Per personal communication with Y. Brehmer (July 29, 2011), additional information that is specific to the training group of Westerberg et al. (2008) along with potential genetic predictors of performance on the training task can be found in Bellander et al. (2011) and Brehmer et al. (2009).

<sup>4</sup> Klingberg et al. (2005) reported an average of forward and backward span performance. Dahlin (2011) separated this data for the control group of Klingberg et al. (2005). Dahlin (2011) additionally included a no-contact control group for any far transfer measures that were not included in the study of Klingberg et al. (2005).

Table 1  
*Relevant Measures of Intelligence and Attention*

Ability of interest	Type of task	Name of task	Description
Fluid intelligence (Gf)	Block design	WASI Performance IQ (Block Design) WPPSI	Rapidly arrange blocks to reproduce a presented pattern
	Figural reasoning	BIS Figural	Figure based analogies, rule discovery, font recognition, figure assembly
	Folding tests	DAT-SR	Choose an abstract figure that may be folded to recreate a 3-D object
	Matrix reasoning	BOMAT DAT-AR Raven's Progressive Matrices TONI WASI Performance IQ (Matrix Test)	A series of abstract figures is arranged in a grid. One piece of the grid is missing. Test-taker must select the missing piece from several options.
	Numerical reasoning	BIS Numerical	Sequence completion, arithmetic, categorization
	Verbal reasoning	BIS Verbal DAT-VR	Verbal analogies, logic, word completion, verbal relationships
Achievement tests	Mathematical	Tempo Test Rekenen WOND	1 min of mixed arithmetic operations Reasoning with mathematical information
	Verbal	Eén Minuut Test WASI Verbal IQ WORD	Read words of increasing difficulty Tests of vocabulary and/or reading comprehension
Attention	Control of attention	PASAT	Series of auditory digits is presented. Test taker must state the sum of the current and previous digit
		Stroop	State the hue in which a word has been printed (e.g., the word <i>BLUE</i> printed in red ink)
		TEA	Counting while distracted, visually searching for items, multitasking
	Sustained attention	Choice reaction time	Rapidly make one of two responses, depending on information that is presented on a computer monitor
		Continuous performance task	Attend to a series of items (e.g., letters), make a response each time a target item is presented
		Go/no-go task	Press a button when one type of stimulus appears on a computer monitor, withhold response when a different type appears

*Note.* BIS = Berlin Intelligence Structure Test; BOMAT = Bochumer Matrizen-Test; DAT = Differential Aptitude Battery; AR = Abstract Reasoning; SR = Spatial Relations; VR = Verbal Reasoning; PASAT = Paced Auditory Serial-Attention Task; TEA = Test of Everyday Attention; TONI = Test of Nonverbal Intelligence; WASI = Wechsler Abbreviated Scales of Intelligence; WOND = Wechsler Objective Number Dimensions; WORD = Wechsler Objective Reading Dimensions; WPPSI = Wechsler Preschool and Primary Scale of Intelligence.

The only study involving children with ADHD or low WM capacity to demonstrate unqualified near transfer was Holmes et al. (2009). Interestingly, this improvement was not only demonstrated with traditional span tasks but also in a task that required children to repeat complex sets of instructions (e.g., Gathercole et al., 2008). As we have stated, children with low WM capacity have difficulty completing such tasks (Engle et al., 1991; Gathercole et al., 2008), and this finding serves as an interesting analogue to real-world behavior.

**Transfer to Gf and achievement tests.** Six studies examined the effect of WM training on Gf (see Table 2). The results are mixed. Both Klingberg et al. (2002) and Klingberg et al. (2005) reported Ravens improvements, relative to nonadaptive control groups. However, neither Holmes et al. (2009) nor Holmes et al. (2010) found any improvement on a different measure of reasoning, following Cogmed training (WASI performance IQ; block design and matrix reasoning subtests; Wechsler, 1999; see Table 1). K. I. E. Dahlin (2011), on the other hand, reported test–retest improvement on Ravens, but this was not significantly different from her control group (ADHD children from Klingberg et al., 2005). Van der Molen et al. (2010) did not find transfer of OddYellow training to Ravens (Raven, 1998).

Five studies included achievement tests in their transfer batteries (see Table 1). Three are directly comparable, due to the inclusion of common tasks. These tests involved numerical computation (Wechsler Objective Number Dimensions [WOND]; Wechsler, 1996) and vocabulary (Wechsler Abbreviated Scale of Intelligence [WASI] Verbal IQ; Wechsler, 1999; Wechsler Objective Reading Dimensions [WORD]; Wechsler, 1993). Holmes et al. (2009) did not find transfer to WASI Verbal IQ, WORD, or WOND. Additionally, Holmes et al. (2010) did not find transfer to WASI Verbal IQ. Alloway (in press), on the other hand, reports transfer to WASI Verbal IQ and WOND, but, as with her WM results, this was relative to a control group who performed numerically worse at posttest.

K. I. E. Dahlin (2011) included tests of reading comprehension, recognition of spelling errors and phonological reading ability.<sup>5</sup>

<sup>5</sup> Reading comprehension tasks were drawn from the Progress in International Reading Literacy Study (<http://nces.ed.gov/surveys/PRLS>) and the International Association for the Evaluation of Educational Achievement ([http://www.iea.nl/reading\\_literacy.html](http://www.iea.nl/reading_literacy.html)). Other tasks were developed in house.

Table 2  
Transfer Results for Studies With Children

Training program (by group characteristic)	Authors	Control group	Near					Far						
			STM	WMC			ADHD Obj.	ADHD Subj.	n	Age in years M (SD)				
				BS	CS	Gf					Ach.	Attn.		
Children with ADHD and/or low WMC														
Cogmed (ADHD)	Beck et al. (2010)	No contact											51	11.75 (32.59)
Cogmed (ADHD)	Gibson et al. (2011)	None											37	12.59 (1.21)
Cogmed (ADHD)	Holmes et al. (2010)	None											25	9.75 (.92)
Cogmed (ADHD)	Klingberg et al. (2002)	Nonadaptive task											14	11.20 (2.55)
Cogmed (ADHD)	Klingberg et al. (2005)	Nonadaptive task											44	9.80 (1.30)
Cogmed (ADHD)	Mezzacappa & Buchner (2010)	None											8	8.75 (.89)
Cogmed (cochlear implants)	Kronenberger et al. (2011)	None											9	10.20 (2.2)
Cogmed (low WMC)	Holmes et al. (2009)	Nonadaptive task											42	8-11 <sup>c</sup>
Cogmed (special education)	K. I. E. Dahlin (2011)	Klingberg et al. (2005)											57	10.71 (1.09)
JungleMemory (learning disability)	Alloway (in press)	Learning support											15	13.00 (.78)
OddYellow (borderline IQ)	Van der Molen et al. (2010)	Response time task											93	15.21 (.69)
Typically developing children														
Cogmed	Bergman Nutley et al. (2011)	Nonadaptive task											101	4.30 (.25)
Cogmed	Shavelson et al. (2008) <sup>b</sup>	Nonadaptive task											37	13.50 (.70)
Cogmed	Thorell et al. (2009)	Computer games											62	4.70 (.43)
n-back	Jaeggi et al. (2011)	Knowledge training											62	9.03 (1.49)
Running span <sup>a</sup>	Zhao et al. (2011)	Computer games											33	9.76 (.61)

Note. ✓ = significant transfer; ? = mixed transfer; dash = no transfer; DNR = did not report; STM = short-term memory; WMC = working memory capacity; BS = backward span; CS = complex span; Gf = general fluid intelligence; Ach. = achievement; Attn. = attention; ADHD = attention-deficit/hyperactivity disorder; Obj. = objective measurement; Subj. = subjective measurement; n = number of participants included in posttest session.

<sup>a</sup> Training task did not adapt to performance. <sup>b</sup> Study not published in a peer reviewed journal. <sup>c</sup> Data unavailable; age range substituted.

Results were mixed. Children improved on the reading comprehension test, but not on the other two.

Finally, Van der Molen et al. (2010) included measures of arithmetic (Tempo Test Rekenen; de Vos, 1992; see Table 1) and reading ability (Eén Minuut Test; Brus & Voeten, 1973; see Table 1). No immediate training effects were apparent.

**Transfer to attention.** Four studies used the Stroop task to measure transfer to attention. Klingberg et al. (2002) reported transfer on this task relative to control participants, and Klingberg et al. (2005) replicated this finding. Van der Molen et al. (2010) did not find transfer of OddYellow training to Stroop performance. This inconsistency may be attributable to differences in training programs (Cogmed vs. OddYellow), or populations from which participants were drawn (ADHD vs. lower IQ). Relevant to this question, K. I. E. Dahlin (2011) used Cogmed training with special education students and did not find transfer to Stroop performance. This suggests the difference may be population specific.

In reference to the Stroop tasks used by Klingberg et al. (2005, 2002), a later article (Klingberg, 2010) states that “control congruent trials were not included” (p. 319). This is a point on which the original reports were vague, and it is consequential. As we have noted, when congruent trials are omitted, WM capacity does not predict performance on the Stroop task (Hutchison, 2007; Kane & Engle, 2003). The exclusion of congruent trials from the Stroop tasks of Klingberg et al. (2002) and Klingberg et al. (2005) implies that the reported performance improvements are not readily explained by increases in WM capacity.

Of further concern, research in this area has thus far relied heavily on the Stroop task. Future studies should employ a variety of tasks that converge on the *attention construct* (e.g., antisaccade, Hallett, 1978; flanker, Eriksen & Eriksen, 1974). This would increase confidence that WM training improves attentional control in children with ADHD/low WM capacity.

**ADHD-related symptoms.** Klingberg et al. (2002) and Klingberg et al. (2005) employed objective measures of ADHD-specific symptoms. In both studies, head movements (i.e., hyperactive behaviors) were recorded while the child completed a continuous performance task (see Table 1). Although Klingberg et al. (2002) found a significant training-related reduction in movements (relative to an active control group), this did not replicate in the double-blind study of Klingberg et al. (2005). Klingberg et al. (2002) also included a choice reaction time task (see Table 1) in their far transfer battery. By the account of Westerberg et al. (2004), performance on this task should reflect the ability to sustain attention, and improvement would thus signal a decrease in inattentive symptoms. However, posttest performance revealed equivocal evidence of transfer.

Gibson et al. (2011) took a different approach by measuring changes with a free recall task (e.g., 12 words are presented, test taker tries to remember as many as possible). Performance on free recall tasks can be divided into two components: primary and secondary memory (cf. Craik & Birtwistle, 1971; Tulving & Colotla, 1970). Primary memory refers to recall of the final items on the list and is conceptually similar to STM. Secondary memory refers to recall of items from early in the list and is conceptually similar to long-term memory. A separate study by Gibson, Gondoli, Flies, Dobrzanski, & Unsworth (2009) found that children with ADHD are specifically deficient on the secondary memory portion of free recall. However, while Gibson et al. (2011) found

that Cogmed training improved primary memory performance, retrieval from secondary memory was unaffected.

Four studies included subjective reports of ADHD symptoms in the form of teacher- and/or parent-provided ratings on forms such as the Conners Rating Scale (Conners, 2001) and the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; American Psychiatric Association, 2000). These inventories allow researchers to judge the presence of symptoms such as inattention and hyperactive behaviors outside of the laboratory. A serious problem is that raters were rarely blind to condition assignment, and a predictable pattern emerged. When raters were aware that children were receiving WM training, they reported behavioral improvements (i.e., teachers in Mezzacappa & Buckner, 2010; parents in Beck et al., 2010; teachers and parents in Gibson et al., 2011). When raters were blind to condition assignment, they did not report behavioral changes (i.e., teachers in Beck et al., 2010; teachers in Klingberg et al., 2005). One exception to this trend was the parents in Klingberg et al. (2005), who, despite being blind to condition assignment, reported training-group-specific decreases in inattention and hyperactivity. Additionally, the nonblind teachers in Gibson et al. (2011) reported decreased inattention but not hyperactivity. Regardless, the majority of subjective data conforms to predictions that could be made on the basis of expectation of outcome alone (e.g., M. Conway & Ross, 1984; DeLoache et al., 2010; Greenwald et al., 1991). When raters know that children are receiving treatment, training-related changes are perceived. When raters are blind, training-related changes are not perceived.

**Working memory training and typically developing children.** The bottom portion of Table 2 details five studies that have been conducted with typically developing children. These studies were not limited to simple/complex span tasks but also included training on *n*-back and running span.

**Transfer to WM capacity.** To date, studies that involve typically developing children have produced mixed near transfer results. In a study focused on middle school children, Shavelson et al. (2008) included two complex span tasks. The training group did not show posttest improvements on either task, relative to an active control group. Bergman Nutley et al. (2011), on the other hand, did find transfer of training to a complex span task for younger children.

The training program employed by Thorell, Lindqvist, Bergman, Bohlin, and Klingberg (2009) was restricted to visuo-spatial tasks. Thus, it is potentially interesting that trained children increased their simple span scores on a verbal task. This finding, however, did not replicate in a later study (Bergman Nutley et al., 2011).

**Transfer to Gf.** No study involving training on span tasks (see Table 2) reported significant transfer to Gf. Shavelson et al. (2008) measured Gf using Ravens, while Thorell et al. (2009) used a block design task (Wechsler Preschool and Primary Scale of Intelligence-Revised; Wechsler, 1995). Bergman Nutley et al. (2011) included both Ravens (three versions) and block-design tasks (Wechsler Preschool and Primary Scale of Intelligence, 3rd ed.; Wechsler, 2004).

Other training methods have reported greater success. Although Jaeggi et al. (2011) found no overall transfer of *n*-back training to a composite Gf measure (Ravens and the Test of Nonverbal Intelligence; Brown, Sherbenou, & Johnsen, 1997), a post hoc analysis revealed that children who made large progress on the

training task also improved on the Gf measure, while the small-progress group did not. Zhao, Wang, Liu, and Zhou (2011) trained children on a nonadaptive running span task. These researchers found transfer to Ravens and further reported a correlation between improvement on the training task and far transfer (.54).

Both Jaeggi et al., 2011 and Zhao et al., 2011 found strong associations between progress in training and far transfer. Thus, further research with both types of training is warranted. However, far transfer cannot be readily attributed to increased WM capacity, as neither study reported near transfer tasks. Of note, an earlier report regarding the data of Jaeggi et al. (2011; Buschkuhl, Jaeggi, Jonides, & Shah, 2010) indicated that *n*-back and a continuous performance task (see Table 1) were included in the transfer battery. These results were not included in the published version but would have allowed for a clearer understanding of the generality of their post hoc analysis (i.e., did improvements on the training task also predict near transfer to WM and far transfer to attention?).

**Transfer to attention.** Thorell et al. (2009) examined the effect of adaptive WM training on the attention of preschoolers, via several measures. Dependent variables included errors on a Stroop-like task (i.e., name the opposite: day/night and boy/girl), omissions and incorrect responses on a go/no-go task (see Table 1) and omissions on a continuous performance task (see Table 1). Relative to control children, trained children showed a decrease in the number of correct responses that were omitted in the go/no-go and continuous performance tasks. This may provide evidence that training on an adaptive WM task improves vigilance. However, as with studies involving children who have a presumed WM deficiency, the evidence of transfer to attention is limited in scope and therefore requires broader study before a confident statement can be made.

### Working Memory Training with Young Adults

Table 3 summarizes 13 WM training experiments<sup>6</sup> that have been conducted with young adults. Due to differences in both results and theoretical perspectives, the studies have been organized by type of training procedure (i.e., simple/complex span task, *n*-back, running span), and further discussion is similarly organized. The training program used by Schmiedek et al. (2010) was not strictly based on WM and is therefore difficult to classify. However, this is one of the few studies to attempt to measure transfer to latent abilities; it is discussed in detail in later sections.

As with studies involving children, control groups are a concern (see Table 3). Nine experiments used no-contact control groups; one did not include a control group; and Klingberg et al. (2002) compared their healthy adult participants to an ADHD-diagnosed control group from a separate experiment (children from Experiment 1 of Klingberg et al., 2002). Thus, with little control for Hawthorne/placebo effects (e.g., Klingberg, 2010; Shipstead et al., 2010), Table 3 provides an optimistic view of the effect of WM training on young adults.

**Preview of results.** There is little evidence that training programs that are based on simple/complex span tasks (top of Table 3) change the cognitive abilities of young adults. *n*-back training (middle of Table 3), on the other hand, has shown promise in terms of transfer to Gf. However, increased WM capacity does not readily explain this effect, and an alternate mechanism of far

transfer has yet to be identified. Finally, E. Dahlin and colleagues (E. Dahlin, Nyberg, et al., 2008; E. Dahlin, Stigsdotter Neely et al., 2008) propose that transfer of training is limited to tasks that tap the same cognitive functions as the method of training (i.e., there is no general effect of training). Direct evidence is preliminary.

**Simple/complex span task training and young adults.** Seven studies that trained young adults on simple/complex span tasks are included in the top section of Table 3. Of note, Chein and Morrison (2010) was the only study to train participants on adaptive complex span tasks, while the program of Colom, Quiroga, et al. (2010) involved only three sessions of practice on nonadaptive tasks.

**Transfer to WM capacity.** There is little evidence that simple/complex span training increases WM capacity in young adults (see Table 3). Most studies have used STM tasks as their measures of near transfer. Chein and Morrison (2010) reported significant training-related improvements on complex span performance. However, the tasks in the near transfer battery were nonadaptive versions of the same tasks on which participants had trained. As such, this transfer effect can be readily attributed to practice.

**Transfer to Gf.** Evidence of transfer to Gf is similarly sparse. Six experiments measured far transfer of simple/complex span training to Gf tasks (see Table 3), but only two reported significant results.<sup>7</sup> Of note, these two studies had training groups of four or fewer participants (i.e., Klingberg et al., 2002; Olesen et al., 2004), while the experiments that reported a lack of far transfer had training groups of 19 or more participants (i.e., Chein & Morrison, 2010; Westerberg et al., 2008). Thus, an effect of simple/complex span training on the reasoning abilities of young adults has yet to be demonstrated in a large-scale study.

**Transfer to attention.** At a glance, Table 3 implies that transfer to attention is among the most promising avenues for simple/complex span training. Of the six experiments that included attention tasks, five report some evidence of transfer.

There are, however, several concerns. Due to choice of control group (i.e., ADHD-diagnosed children; Table 1), the results of Klingberg et al. (2002) are best viewed as a test-retest effect. Moreover, Klingberg (2010) reported that Stroop tasks in Klingberg et al. (2002) and Olesen et al. (2004) did not include congruent trials. Thus, as with studies involving children, training-related improvements cannot be readily attributed to increased WM capacity (cf. Hutchison, 2007; Kane & Engle, 2003). This concern is reinforced by Experiment 2 of Olesen et al. (2004), which found far transfer to performance of the Stroop task but equivocal evidence of near transfer to two simple span tasks. This further suggests that something other than increased WM capacity is at the root of far transfer to Stroop performance. Finally, the results of Chein and Morrison (2010; 50% congruent) indicate that the effect of adaptive complex span training on Stroop performance is, at best, weak. While one-tailed *t* tests revealed that the training and no-contact groups differed in their posttest Stroop performance, the relevant interaction (Testing Session  $\times$  Group) was small ( $\eta_p^2 = .056$ ) and failed to reach statistical significance.

<sup>6</sup> As reported by Dahlin, Nyberg, et al. (2008, p. 772), Dahlin, Stigsdotter Neely, et al., is a subset of the data in Dahlin, Nyberg, et al. (2008).

<sup>7</sup> The online supplement of McNab et al. (2009) reported that participants were tested on Ravens, but the results were not reported.

Table 3  
Transfer Results for Studies With Young Adults

Training program	Authors	Control group	Near		Far		<i>n</i>	Age in years <i>M</i> ( <i>SD</i> )	
			WMC			Gf			Attn
			STM	CS	<i>n</i> -back				
Simple/complex span training									
Complex span	Chein & Morrison (2010)	No contact		✓	—	?	42	20.10 (1.74)	
Complex/simple span <sup>a</sup>	Colom, Quiroga, et al. (2010)	Speed/attention tasks			—		288	20.10 (3.40)	
Cogmed	Klingberg et al. (2002; E2)	Children w/ADHD	✓		✓	✓	4	23.50 (3.40)	
Cogmed	McNab et al. (2009)	None	✓		DNR	DNR	13	20–28 <sup>c</sup>	
Cogmed	Olesen et al. (2004; E1)	No contact	✓		✓	✓	3 <sup>d</sup>	20–23 <sup>c</sup>	
Cogmed	Olesen et al. (2004; E2)	No contact	—			✓	8 <sup>d</sup>	29.30 (2.1)	
Cogmed	Westerberg et al. (2008) <sup>b</sup>	Nonadaptive task	✓		—	?	55	26.23 (2.83) <sup>c</sup>	
<i>n</i> -back training									
Dual <i>n</i> -back	Jaeggi et al. (2008)	No contact	✓	—		✓	69	25.6 (3.30)	
Dual/single <i>n</i> -back	Jaeggi, Studer-Luethi, et al. (2010)	No contact		—	✓	?	89	19.4 (1.50)	
Single <i>n</i> -back <sup>a</sup>	Li et al. (2008)	No contact		—	✓		46	25.95 (2.57)	
Dual <i>n</i> -back	Seidler et al. (2010)	Knowledge training		✓	✓	—	56	21.4 (4.82)	
Running span training									
Running span	E. Dahlin, Nyberg, et al. (2008)	No contact	—	—	✓	—	28	23.59 (2.62)	
Running span	E. Dahlin, Stigsdotter Neely, et al. (2008, E1)	No contact			✓	—	22	23.59 (2.54)	
Other WM training									
Various tasks <sup>a</sup>	Schmiedek et al. (2010)	No contact		—	✓	✓	145	25.47 (2.64)	

Note. ✓ = significant transfer; ? = mixed transfer; dash = no transfer; DNR = did not report; STM = short-term memory; WMC = working memory capacity; CS = complex span task; Gf = general fluid intelligence; Attn. = attention; *n* = number of participants included in posttest session.

<sup>a</sup> Training task did not adapt to performance. <sup>b</sup> Study not published in a peer-reviewed journal. <sup>c</sup> Data unavailable; age range substituted. <sup>d</sup> Unclear whether same control participants used in both experiments, therefore only trained participants are listed. <sup>e</sup> Trained group only, as reported in Brehmer et al. (2009).

The largest concern, however, regards the unpublished study of Westerberg et al. (2008). In this study, 55 young adults performed the Stroop, the Paced Auditory Serial-Addition Task (PASAT; see Table 1; Gronwall, 1977), and a choice reaction time task (see Table 1). Although participants improved on the PASAT and became less variable in their reaction times, they did not improve on the Stroop task. Thus, the results of Westerberg et al. (2008), which was the only study to use an active control group, contradict the findings of other studies.

***n*-back training with young adults.** Jaeggi et al. (2008); Seidler et al. (2010), and one of two conditions in Jaeggi, Studer-Luethi, et al. (2010) use a “dual *n*-back” task. The dual *n*-back provides a constant challenge to the limits of WM by requiring participants attend to simultaneously presented visual and auditory stimuli and respond when either matches a stimulus that was presented *n* items ago.

**Transfer to WM capacity.** The only *n*-back training study to show transfer to complex span was the unpublished report of Seidler et al. (2010). Jaeggi and colleagues (Jaeggi et al., 2008; Jaeggi, Studer-Luethi, et al., 2010) twice failed to detect near transfer from adaptive *n*-back training to complex span tasks. A similar lack of transfer between *n*-back and complex span has been found in studies in which participants trained on nonadaptive *n*-back tasks (Li et al., 2008; Schmiedek et al., 2010). It should be noted that several studies have concluded that *n*-back and complex span tasks tap related but separable aspects of WM capacity (Ilkowska, 2011; Jaeggi, Buschkuhl, et al., 2010; Kane, Conway, Miura, & Colflesh, 2007; Schmiedek et al., 2009). This may explain the lack of near transfer; however, a new problem arises:

Far transfer of adaptive *n*-back training cannot be easily attributed to a general increase in WM capacity.

**Transfer to Gf.** Jaeggi and associates (Jaeggi et al., 2008; Jaeggi, Studer-Luethi, et al., 2010) have twice demonstrated far transfer of training to matrix reasoning tasks (see Table 3). In order to reconcile the finding of far transfer with the absence of near transfer Jaeggi et al. (2008) initially proposed that the divided attention aspect of the dual *n*-back might have strengthened executive function, which, in turn, led to an increase in Gf.

This hypothesis was subsequently tested by Jaeggi, Studer-Luethi, et al. (2010), who included two training conditions: dual *n*-back and single *n*-back (visuo-spatial component only). For the single *n*-back condition, far transfer to Ravens (Raven, 1990) and BOMAT (see Table 1) was found. Thus, the divided-attention hypothesis was falsified. More problematic, the dual *n*-back group showed transfer to Ravens but not to BOMAT.<sup>8</sup> BOMAT was the Gf task used to demonstrate transfer in the original study of Jaeggi et al. (2008). Thus, this second study (Jaeggi, Studer-Luethi, et al., 2010) may also be interpreted as a failure to replicate an earlier finding from the same research group.

These complications aside, Jaeggi and colleagues (Jaeggi et al., 2008; Jaeggi, Studer-Luethi, et al., 2010; Seidler et al., 2010) have twice found transfer of *n*-back training to Gf tasks. Thus, for adults, *n*-back training has shown promise relative to simple/

<sup>8</sup> This analysis was not included in the original report of Jaeggi, Studer-Luethi, et al. (2010) but was confirmed via personal communication with S. M. Jaeggi (September 15, 2010).

complex span training. However, future challenges include (a) identifying the near transfer mechanism of  $n$ -back training effects, (b) demonstrating transfer to nonmatrix reasoning tasks (cf. Moody, 2009), and (c) demonstrating transfer relative to an active control group.

**Running span training with young adults.** E. Dahlin and associates (E. Dahlin, Nyberg, et al., 2008; E. Dahlin, Stigsdotter Neely, et al., 2008) hypothesized that transfer effects should be function specific. For instance, if a training task involves updating immediate memory, then transfer should only occur for tasks that require memory updating. E. Dahlin, Stigsdotter Neely, et al. (2008) demonstrated via fMRI that the running span and  $n$ -back elicit common activation in the striatum, which is a potential gateway for WM updating (McNab & Klingberg, 2008; see also Awh & Vogel, 2008). It was thus assumed that both running span and  $n$ -back require memory updating, and training on one should transfer to the other.

Consistent with this hypothesis, E. Dahlin, Stigsdotter Neely, et al. (2008) found transfer of running span training to the  $n$ -back and posttraining increases in striatal activation that were common to both tasks. Critically, no transfer was found for the Stroop task, which did not elicit striatal activation. Ostensibly consistent with their function-specific hypothesis, transfer to complex span was also absent (E. Dahlin, Nyberg, et al., 2008); however, fMRI data were not available for this task.

Although this hypothesis is interesting, complex span and updating tasks (in particular the running span) are typically found to be highly related and are often argued to reflect the same cognitive mechanisms (Broadway & Engle, 2010; Miyake et al., 2000; see also Schmiedek et al., 2009). Therefore, it is unclear why complex span performance should not benefit from running span training.

Of particular concern, the running span and  $n$ -back share a great deal of feature overlap (see Figure 2). Both involve attending to a series of serially presented items, and keeping track of only the last few. Thus, task-relevant practice provides a plausible alternate account of the limited transfer findings reported by E. Dahlin, Nyberg, et al. (2008) and E. Dahlin, Stigsdotter Neely et al. (2008).

Future studies will need to eliminate this possibility by demonstrating function-specific transfer to a broader range of tasks.

## Working Memory Training and Older Adults

**Transfer to WM capacity.** Table 4 indicates that older adults typically show near transfer to WM tasks that match the method of training, but not to WM tasks that are dissimilar. One exception is the study of Schmiedek et al. (2010), who report that, after training on a battery that included the  $n$ -back, older adults did not show transfer to an  $n$ -back task but did show transfer to one of three complex span tasks.

E. Dahlin, Stigsdotter Neely, et al. (2008) did not find transfer of running span training to  $n$ -back for older adults. This is noteworthy because, in contrast to younger participants (who did show transfer), older participants did not show pretraining striatal activation when performing the running span. Training did increase activation in this area; however, this was only apparent when participants were performing the running span. Striatal activation remained absent when the  $n$ -back was performed. This was taken as an indication that the extent of transfer is limited by age-related brain deficiencies.

**Transfer to Gf.** Four experiments tested participants on Ravens (see Table 4), but only Schmiedek et al. (2010) reported significant transfer. This finding, however, is qualified by two further results. First, no transfer occurred for three other reasoning tasks (subscales of the Berlin Intelligence Structure Test; Jäger, Süß, & Beauducel, 1997; see Table 1). Second, due to the large participant sample and use of multiple tests, Schmiedek et al. (2010) were able to test transfer of training at the latent-factor level (e.g., Figures 3 and 4). However, transfer to latent Gf was absent, thus suggesting the Ravens improvements were task specific.

**Transfer to attention.** Richmond, Morrison, Chein, and Olson (2011) did not find transfer to a battery of attention tasks that involved counting and/or visually searching for specific stimuli (Test of Everyday Attention; Robertson, Ward, Ridgeway, &

Table 4  
*Transfer Results for Studies With Older Adults*

Training program	Authors	Control group	WMC						Age in years <i>M (SD)</i>	
			STM	CS	$n$ -back	Gf	Attn.	LTM		
Simple/complex span training										
Complex/simple span	Buschkuehl et al. (2008)	Cardio-exercise	?					—	32	80.00 (3.30)
Complex span	Richmond et al. (2011)	Trivia	—	✓		—	—	?	40	66.00 (5.46)
Cogmed	Brehmer et al. (2011)	Nonadaptive task	?			—	?	—	24	60–70 <sup>b</sup>
$n$ -back training										
$n$ -back <sup>a</sup>	Li et al. (2008)	No contact		—	✓				41	73.91 (2.8)
Running span training										
Running span	E. Dahlin, Nyberg, et al. (2008)	No contact		—	—	—		—	27	68.31 (1.69)
Running span	E. Dahlin, Stigsdotter Neely, et al. (2008, E2)	No contact			—				19	68.31 (1.84)
Other WM training										
Various tasks <sup>a</sup>	Schmiedek et al. (2010)	No contact		?	—	?		?	142	71.09 (4.07)

Note. ✓ = significant transfer; ? = mixed transfer; dash = no transfer; STM = short-term memory; WMC = working memory capacity; CS = complex span task; Gf = general fluid intelligence; Attn. = attention; LTM = long-term memory;  $n$  = number of participants included in posttest session.

<sup>a</sup> Training task did not adapt to performance. <sup>b</sup> Data unavailable; age range substituted.



Nimmo-Smith, 1994; see Table 1). Brehmer et al. (2011), on the other hand, reported limited transfer that mirrored the younger participants of Westerberg et al. (2008)<sup>9</sup>: Older participants improved their performance on PASAT and a choice reaction time task (see Table 1), but transfer to Stroop performance was absent.

**Retrieval from long-term memory.** Age-related declines in WM capacity can largely account for age-related declines in memory retrieval (McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010; Park et al., 1996; see also Salthouse & Babcock, 1991). Several studies thus examined performance on long-term memory tasks. However, evidence of transfer was limited.

Many studies focused on delayed free recall tests, in which participants attempt to recall a long list of items after a delay of several minutes. Neither Buschkuhl et al. (2008) nor E. Dahlin, Stigsdotter Neely, et al. (2008) found transfer (though the training group in Buschkuhl et al., 2008, did show a test–retest effect that was absent for controls). Brehmer et al. (2011) reported only marginal improvement. Richmond et al. (2011) found that participants were less likely to repeat items during recall, but the overall number of items recalled did not increase.

Two studies tested transfer using paired associates tasks. These tasks present a series of word pairs, and later require test takers to recall one word when the other is shown. The older participants of Schmiedek et al. (2010) improved on this task; however, transfer was not apparent in the results of E. Dahlin, Stigsdotter Neely, et al. (2008).

### Duration of Transfer

Table 5 presents transfer results for studies that included follow-up sessions. For ease of viewing, tasks that did not show transfer at either posttest are omitted. Attention tasks are not reported on Table 5 but can be summarized: The results of Westerberg et al. (2008) remained stable at posttest, while Klingberg et al. (2005) did not find long-term transfer to Stroop. This was not due to a decline in the training group's performance; rather, the control group improved at the follow-up session. Additionally, the subjective ADHD scores of Beck et al. (2010) and Klingberg et al. (2005) remained stable.

**Transfer to WM capacity.** Near transfer results were mostly stable. Notable exceptions include Van der Molen et al. (2010) and Kronenberger et al. (2011). At the first posttest, Van der Molen et al. found transfer to verbal STM but not visuo-spatial STM. At the follow-up session these results reversed. Additionally, Van der Molen et al. reported that a group that was performing a low-difficulty, nonadaptive version of OddYellow showed signs of transfer to WM that were not apparent at the original posttest. However, this did not occur for the group that performed rigorous training. Kronenberger et al. (2011), who were conducting a pilot study on children with cochlear implants, found that evidence of near transfer had disappeared by the 6-month follow-up. Despite this, far transfer to a sentence repetition task remained. These sentences were drawn from a limited pool, and thus test–retest effects cannot be ruled out (particularly in the absence of a control group). However, the researchers argue that the 6-month delay between sessions contradicts this interpretation.

Finally, near transfer was absent for the older adults who were included in the study of Buschkuhl et al. (2008). They did,

however, show improvements in verbal free recall that were not apparent at the first posttest.

**Transfer to Gf and achievement.** It is reasonable to assume that transfer to tests of academic attainment require time before the effect of training will be seen. Holmes et al. (2009) reported that at the 6-month follow-up, children who received WM training improved their performance on the mathematical reasoning subtest of WOND (see Table 1), but not on verbal tests (WASI Verbal IQ; WORD; see Table 1). Similarly, Van der Molen et al. (2010) reported that children who were trained on either low-difficulty or adaptive OddYellow improved on an arithmetic test (Tempo Test Rekenen; see Table 1) but not on a reading test (Eén Minuut Test; see Table 1). In addition to these findings, K. I. E. Dahlin (2011) reported that improved reading comprehension remained stable across the 6- to 7-month delay.

The long-term improvement reported by Holmes et al. (2009), however, was within group. The control group was not included in the follow-up. Thus, test–retest effects cannot be ruled out (also maturation effects; Campbell & Stanley, 1963). This alternate account is strengthened by the results of Klingberg et al. (2005) and Jaeggi et al. (2011). In both studies, training-related improvements on Gf tasks were diminished or absent at the follow-up posttest. These effects, however, were not driven by a regression of training group scores but rather by improved control-group scores over the course of three testing sessions. Nonetheless, the long-term effect of WM training on scholastic achievement tests is an understudied area that requires future research.

### Attributing Training Effects to Latent Change

Three of the reviewed studies have attempted to demonstrate transfer of training to latent abilities rather than to single tasks. Two (Bergman Nutley et al., 2011; Schmiedek et al., 2010) have found evidence of latent change; however, neither instance was attributable to WM. A third study (Colom, Quiroga, et al., 2010) found large improvements on several measures of Gf but no evidence of latent change.

**Schmiedek et al. (2010).** Schmiedek et al. (2010) did not specifically train participants on a WM task. Rather, participants completed 100 sessions of practice on several tasks (e.g., *n*-back, episodic memory, reaction time). Using a latent change model (based on McArdle & Prindle, 2008), Schmiedek et al. (2010; see Table 3) report that younger adults showed a small, but statistically significant, increase in Gf as defined by Ravens and the three Berlin Intelligence Structure tests listed in Table 1. Thus, unlike transfer of training to single tasks, this finding indicates that cognitive training may change latent abilities. However, while near transfer was found for WM tasks that were similar to tasks found in the training program (including *n*-back), near transfer to untrained complex span tasks was not found. Thus, while Schmiedek et al. (2010) did find transfer to Gf (for younger adults), this effect cannot be specifically attributed to time spent on WM training or an increase in WM capacity (both of which are assumed to be critical components of a successful training program; e.g., Klingberg, 2010).

<sup>9</sup> Per personal communication with Y. Brehmer (July 20, 2011), participants reported in Brehmer et al. (2011) were older participants from the unpublished study of Westerberg et al. (2008).

Table 5  
Long-Term Transfer

Training program	Group	Authors	Months since training	STM	WMC			Gf	Ach.
					BS	CS	<i>n</i> -back		
Children									
Cogmed	ADHD	Holmes et al. (2010)	6	?		✓			
Cogmed	ADHD	Klingberg et al. (2005)	3	✓				—	
Cogmed	Cochlear implants	Kronenberger et al. (2011)	6	—	—				
Cogmed	Low WMC	Holmes et al. (2009)	6	?		✓			?
Cogmed	Special education	K. I. E. Dahlin (2011)	6–7	?	?				✓
OddYellow	Borderline IQ	Van der Molen et al. (2010)	2.5	?		?			?
<i>n</i> -back	Typically developing	Jaeggi et al. (2011)	3					?	
Younger adults									
Cogmed		Westerberg et al. (2008)	3	✓					
<i>n</i> -back <sup>a</sup>		Li et al. (2008)	3					✓	
Running span		E. Dahlin, Nyberg et al. (2008)	18					✓	
Older adults									
Complex/simple span		Buschkuhl et al. (2008)	12	—					
<i>n</i> -back <sup>a</sup>		Li et al. (2008)	3					✓	
Running span		E. Dahlin, Nyberg, et al. (2008)	18						

Note. ✓ = significant transfer remained; ? = mixed transfer; dash = transfer regressed; STM = short-term memory; WMC = working memory capacity; BS = backward span; CS = complex span; Gf = general fluid intelligence; Ach. = achievement.

<sup>a</sup> Training task did not adapt to performance.

**Bergman Nutley et al. (2011).** As indicated in Table 2, Bergman Nutley et al. (2011) did not find transfer of WM training to Gf. This experiment did, however, include two other conditions in which children were trained on adaptive, nonverbal reasoning tasks. These tasks involved pattern completion, logical progression, and classification. Difficulty was altered through increasing sequence lengths and the number of relevant dimensions in the puzzle (e.g., size, shape, etc.). At posttest, children who had been trained using this procedure showed increases (relative to a non-adaptive WM control group) on a factor composed of three variations of Ravens along with WPPSI Block Design (Wechsler, 2004; see Table 1).

This may prove to be an important finding. However, the relationship of the training program to the transfer tasks is a concern. Children were provided with extensive practice at finding patterns among nonverbal items, a skill that is important to matrix reasoning and block design tasks but may not be applicable to other contexts. Thus, the interesting results of Bergman Nutley et al. (2011) would be made more persuasive by follow-up studies that demonstrate transfer even when Gf is more broadly defined.

**Colom, Quiroga, et al. (2010).** A recent study by Colom, Quiroga, et al. (2010) highlights our concern regarding the use of single tasks as measure of abilities. In this study, participants received three sessions of practice on either a battery of nonadaptive WM tasks or a battery of nonadaptive processing-speed tasks and attention tasks. Transfer was tested using Ravens and subcomponents of the Differential Aptitude Test battery (DAT; abstract reasoning [AR], verbal reasoning [VR], and spatial relations [SR]; see Table 1). Large test–retest effects were found on three of four reasoning tasks (Ravens, DAT-VR, DAT-SR). However, these effects were present in both the WM and the non-WM groups.

Colom, Quiroga, et al. (2010), nonetheless, tested whether these test–retest improvements could be attributed to increases in Gf. This was accomplished using Jensen’s (1998) method of correlated

vectors. In this technique, the factor loadings of a variety of tests (e.g., loadings of Gf) are first obtained using untrained participants. Some tests should load highly on the factor of interest, while others should not. If a training program affects this factor (e.g., if Gf increases), then performance improvements should be greatest for the tasks with the highest pretest loadings and smallest for the tasks with the lowest loadings.

Colom, Quiroga, et al. (2010) could not attribute the performance improvements to increased Gf. Of the tasks in their battery, the one that had the highest *g*-loading (DAT-AR; Bennett, Seashore, & Wesman, 1990; see Table 1) was also the only reasoning task on which participants did not show a training effect. Thus, as we have argued, improved performance on individual tests is not sufficient to conclude that an underlying ability has changed.

## Summary and Future Questions

### Do Training Programs Increase WM Capacity?

We began by stating that rather than attempting to judge the efficacy of WM training based upon the presence or absence of far transfer results (e.g., Klingberg, 2010; Morrison & Chein, 2011; Shipstead et al., 2010), our conclusions would focus on how well the fundamental assumptions of the training literature have thus far been supported. The first assumption is that WM training increases WM capacity. Verifying this claim should be a prominent goal of the literature, as increased WM capacity is the mechanism through which far transfer is hypothesized to occur. The literature review reveals two prominent concerns: (a) studies that test near transfer using STM tasks and (b) near transfer tasks that closely resemble the method of training.

Examining studies that trained participants with simple/complex span tasks (see Tables 2, 3, and 4), evidence of near transfer is often demonstrated via simple span. This is particularly true for

studies involving adults. Only Chein and Morrison (2010) and Richmond et al. (2011) have attempted to demonstrate transfer via complex span tasks. Although both studies reported near transfer, Chein and Morrison's effect was obtained using the task on which participants had trained. Studies involving children have been more consistent in using WM tasks as tests of near transfer. However, while many of these studies reported transfer to WM, several did not include adequate control groups (K. I. E. Dahlin, 2011; Holmes et al., 2010; Kronenberger et al., 2011; Mezzacappa & Buckner, 2010) or reported equivocal results (Alloway, in press). Of the Cogmed and OddYellow studies that included control groups (see Table 2), only two (Bergman Nutley et al., 2011; Holmes et al., 2009) demonstrated significant transfer using WM tasks. Thus, there is clear need for future studies to consistently employ validated WM tasks, in conjunction with adequate control groups.

However, a larger concern regards transfer across categories of WM tasks. As we have noted, span tasks are not simply related to one another by WM but also by several task-relevant features (e.g., remembering certain types of information in sequence). We therefore make the simple observation that studies in which participants are trained on simple/complex span tasks have thus far avoided attempting to demonstrate near transfer to a broad variety of WM tasks (e.g., *n*-back or running span). Moreover, studies that trained participants on *n*-back and running span (lower portions of Tables 2, 3, and 4) have often attempted to demonstrate near transfer to complex span tasks but rarely found it.

This latter trend may be interpreted in one of two ways. On one hand, it may provide evidence that transfer from *n*-back or running span tasks does not improve WM in general. Rather, benefits of training are restricted to certain cognitive functions (e.g., E. Dahlin, Nyberg, et al., 2008; E. Dahlin, Stigsdotter Neely, et al., 2008; Jaeggi, Studer-Luethi, et al., 2010). However, a more parsimonious explanation is that near transfer does not represent training of WM, but practice with certain varieties of WM task. For instance, learning (e.g., practice, strategies) that occurs during *n*-back and running span training may simply not apply to complex span tasks.<sup>10</sup> To eliminate this possibility, a goal of future studies must be to demonstrate that near transfer can occur, even when task-specific overlap between training programs and transfer tasks has been minimized.

### Does Working Memory Explain Far Transfer?

Verifying that near transfer effects are robust to task-specific changes will be an important step toward establishing the validity of WM training. However, a second concern is the degree to which far transfer results can be explained by increases in WM capacity. At this point, a direct link has not been established. In studies that have been successful in finding signs of latent change, WM training has either been only one aspect of a larger regimen (Schmiedek et al., 2010), or found to be ineffective (Bergman Nutley et al., 2011). Thus, there is a clear need to identify the source of far transfer results.

**Far transfer to Gf.** For children, 11 of the reviewed studies attempted to measure transfer to Gf. Of the nine that trained children using simple/complex span tasks, only two report transfer to Gf (Klingberg et al., 2005, 2002). This low success ratio may reflect a need to better understand the factors involved in transfer

(e.g., Jaeggi et al., 2011), such as the training potential of a given population (e.g., Klingberg, 2010). However, an unavoidable concern is that for simple/complex span training, transfer to Gf has only been demonstrated using Ravens. Transfer to other reasoning tasks has yet to be demonstrated. It therefore remains parsimonious to assume that, rather than increasing WM capacity, the training program used by Klingberg et al. (2005, 2002) affected something that is important to Ravens but not general to all reasoning tasks (cf. Moody, 2009).

The early results of other WM training methods are promising. In studies with children, Jaeggi et al. (2011) and Zhao et al. (2011) found strong correlations between training task progress and improvements on Gf tasks. However, these results are preliminary, and replication with a wider variety of Gf tasks is required. More important, future studies should attempt to pinpoint the source of these changes by reporting near transfer tasks and attempt to understand why some children seem to benefit from WM training, while others do not.

For adults, there is virtually no evidence that simple/complex span training increases Gf (see Tables 3 and 4). In contrast, training with *n*-back has produced more encouraging results. Two *n*-back studies that tested far transfer to Gf reported significant gains (Jaeggi et al., 2008; Jaeggi, Studer-Luethi, et al., 2010), while one did not. Interestingly, the study that did not find far transfer (Seidler et al., 2010) was also the only study to report near transfer to a WM task other than *n*-back. At this point there is little evidence to indicate that WM is responsible for transfer from *n*-back training to Gf.

Identifying the source of far transfer in the studies of Jaeggi et al. (2008) and Jaeggi, Studer-Luethi, et al. (2010) is further confounded by the use of no-contact control groups. This concern is highlighted by the results of Colom, Quiroga, et al. (2010), who, after only three training sessions, found large test–retest gains on several Gf tasks. Importantly, this experiment is made comparable to Jaeggi, Studer-Luethi, et al. (2010) by the common inclusion of Ravens in their far transfer batteries. To reiterate the findings of Colom, Quiroga, et al. (2010), test–retest improvements occurred regardless of whether participants trained on WM or processing speed tasks, and further analysis via the method of correlated vectors (Jensen, 1998) indicated that the gains did not represent increased Gf.

Despite limited training, Colom, Quiroga, et al. (2010) reported that their WM and perceptual speed groups showed pretest–posttest improvements on Ravens ( $d = 0.46$  and  $.61$ , respectively) that were surprisingly comparable in magnitude to the pretest–posttest improvements shown by the single and dual *n*-back groups of Jaeggi, Studer-Luethi, et al. (2010;  $d = 0.65$  and  $.98$ , respectively) after 20 sessions. Critically, the no-contact control group of Jaeggi, Studer-Luethi, et al. (2010) did not produce even a test–retest effect ( $d = 0.09$ ). Thus, Colom, Quiroga, et al. (2010) were

<sup>10</sup> It is sometimes implied that, because strategies are easily undermined (e.g., Chase & Ericsson, 1982), near transfer cannot be explained by strategy formation (e.g., Klingberg, 2010). This, however, is an assumption that is not put to the test in any of the reviewed articles. Indeed, explicit strategy training can produce effects that are quite similar to those produced by adaptive WM training programs (cf. St. Clair-Thompson, Stevens, Hunt, & Bolder, 2010).

able to produce test–retest effects that were substantially larger than those produced by no-contact groups while using a program that was neither focused on WM nor required prolonged effort. It is clear that, in addition to proposing a mechanism of change, the results of Jaeggi et al. (2008) and Jaeggi, Studer-Luethi, et al. (2010) require replication relative to an active control group. This has been attempted; however, no evidence of transfer to Ravens was found (Seidler et al., 2010).

**Far transfer to attention.** Although several studies have reported training-related improvements in Stroop performance, WM cannot readily account for cases in which congruent trails were excluded (i.e., Klingberg et al., 2005, 2002; Olesen et al., 2004; as reported in Klingberg, 2010). Chein and Morrison (2010) found indications of a small effect of training on performance, but the relevant interaction (against a no-contact control group) was nonsignificant. Although limited transfer was reported by Thorell et al. (2009) and Westerberg et al. (2008; see also Westerberg et al., 2007), the evidence of transfer to attention is sparse. Future studies will need to demonstrate transfer with a greater variety of attention tasks, particularly tasks that have been specifically linked to individual differences in WM capacity.

### Concluding Remarks

An ostensible strength of WM training is that it provides a focused, theoretically motivated method through which broad cognitive change may be stimulated (Klingberg, 2010; Sternberg, 2008). However, contrary to the reports provided at the beginning of this article (and contrary to the claims of commercial providers), the present literature provides insufficient evidence of its efficacy. Our primary concerns regard the need for researchers to (a) include multiple measures of abilities of interest, (b) consistently measure near transfer with valid WM capacity tasks that differ from the method of training, (c) eliminate the use of no-contact control groups, and (d) ensure that when subjective measures of change are used, raters are blind to condition assignment. Until these controls are consistently applied, the meaningfulness of training effects cannot be evaluated.

### References

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30–60. doi:10.1037/0033-2909.131.1.30
- Aiken, L. S., & West, S. G. (1990). Invalidity of true experiments: Self-report pretest biases. *Evaluation Review*, *14*, 374–390. doi:10.1177/0193841X9001400403
- Alloway, T. P. (in press). Can interactive working memory training improve learning? *Journal of Interactive Learning Research*.
- Alloway, T. P., & Alloway, R. G. (2008). *Jungle memory training program*. Edinburgh, United Kingdom: Memosyne.
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliot, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*, 606–621. doi:10.1111/j.1467-8624.2009.01282.x
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*, 1698–1716. doi:10.1111/j.1467-8624.2006.00968.x
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Awh, E., & Vogel, E. K. (2008). The bouncer in the brain. *Nature Neuroscience*, *11*, 5–6. doi:10.1038/nn0108-5
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York, NY: Academic Press. doi:10.1016/S0079-7421(08)60452-1
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612–637. doi:10.1037/0033-2909.128.4.612
- Beck, S. J., Hanson, C. A., Puffenberger, S. S., Benninger, K. L., & Benniger, 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. doi:10.1080/15374416.2010.517162
- Bell, M., Bryson, G., & Wexler, B. E. (2003). Cognitive remediation of working memory deficits: Durability of training effects in severely impaired and less severely impaired schizophrenia. *Acta Psychiatrica Scandinavica*, *108*, 101–109. doi:10.1034/j.1600-0447.2003.00090.x
- Bellander, M., Brehmer, Y., Westerberg, H., Karlson, S., Fürth, D., Bergman, O., . . . Bäckman, L. (2011). Preliminary evidence that allelic variation in the LMX1a gene influences training-related working memory improvement. *Neuropsychologia*, *49*, 1938–1942. doi:10.1016/j.neuropsychologia.2011.03.021
- Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1990). *Differential Aptitude Test* (5th ed.). San Antonio, TX: Psychological Corporation.
- Bergman Nutley, S., Söderqvist, S., Bryde, S., Thorell, L. B., Humphreys, K., & Klingberg, T. (2011). Gains in fluid intelligence after training non-verbal reasoning in 4-year-old children: A controlled, randomized study. *Developmental Science*, *14*, 591–601. doi:10.1111/j.1467-7687.2010.01022.x
- Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Bécic, E. (2005). Training effects on dual-task performance: Are there age-related differences in plasticity of attentional control? *Psychology and Aging*, *20*, 695–709. doi:10.1037/0882-7974.20.4.695
- Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., & Baxter, C. (2011). Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry*, *69*, 260–265. doi:10.1016/j.biopsych.2010.08.017
- Borgaro, S., Pogge, D. L., DeLuca, V. A., Bilginer, L., Stokes, J., & Harvey, P. D. (2003). Convergence of different versions of the continuous performance test: Clinical and scientific implications. *Journal of Clinical and Experimental Neuropsychology*, *25*, 283–292. doi.org/10.1076/jcen.25.2.283.13646. doi:10.1076/jcen.25.2.283.13646
- Brehmer, Y., Riekmann, A., Bellander, M., Westerberg, H., Fischer, H., & Bäckman, L. (2011). Neural correlates of training-related working-memory gains in old age. *NeuroImage*, *58*, 1110–1120. doi:10.1016/j.neuroimage.2011.06.079
- Brehmer, Y., Westerberg, H., Bellander, M., Fürth, D., Karlsson, S., & Bäckman, L. (2009). Working memory plasticity modulated by dopamine transporter genotype. *Neuroscience Letters*, *467*, 117–120. doi:10.1016/j.neulet.2009.10.018
- Broadway, J. M., & Engle, R. W. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods*, *42*, 563–570. doi:10.3758/BRM.42.2.563
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *TONI-3: Test of nonverbal intelligence* (3rd ed.). Austin, TX: Pro-Ed.
- Brus, B. T., & Voeten, M. H. M. (1973). *Eén Minuut Test* [One-Minute Reading Test]. Nijmegen, the Netherlands: Berkhout.
- Bühner, M., König, C. J., Prick, M., & Krumm, S. (2006). Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance. *Human Performance*, *19*, 253–275. doi:10.1207/s15327043hup1903\_4
- Bunting, M. F., Cowan, N., & Saults, J. S. (2006). How does running memory span work? *The Quarterly Journal of Experimental Psychology*, *59*, 1691–1700. doi:10.1080/17470210600848402

- Buschkuhl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Däpp, C., Müller, M., . . . Perrig, W. J. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology and Aging, 23*, 743–753. doi:10.1037/a0014342
- Buschkuhl, M., Jaeggi, S. M., Jonidas, J., & Shah, P. (2010). *Working memory training in typically developing children: Interindividual differences and transfer*. Poster presented at the 51st Annual Meeting of the Psychonomic Society, St. Louis, MO.
- Butterfield, E. C., Wambold, C., & Belmont, J. M. (1973). On the theory and practice of improving short-term memory. *American Journal of Mental Deficiency, 77*, 654–659.
- Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand-McNally.
- Cane, V. R., & Heim, A. W. (1950). The effects of repeated retesting: III. Further experiments and general conclusions. *The Quarterly Journal of Experimental Psychology, 2*, 182–197. doi:10.1080/17470215008416596
- Cannon, R., Lubar, J., Gerke, A., Thornton, K., Hutchens, T., & McCammon, V. (2006). EEG spectral-power and coherence: LORETA neurofeedback training in the anterior cingulate gyrus. *Journal of Neurotherapy, 10*, 5–31. doi:10.1300/J184v10n01\_02
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies*. New York, NY: Cambridge University Press. doi:10.1017/CBO9780511571312
- Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386–404. doi:10.1016/0022-0965(82)90054-6
- Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1–58). New York, NY: Academic Press. doi:10.1016/S0079-7421(08)60546-0
- 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. doi:10.3758/PBR.17.2.193
- Clayton, I. C., Richards, J. C., & Edwards, C. J. (1999). Selective attention in obsessive-compulsive disorder. *Journal of Abnormal Psychology, 108*, 171–175. doi:10.1037/0021-843X.108.1.171
- Cogmed. (n.d.). Benefits. Retrieved from <http://www.cogmed.com/benefits>
- Cogmed. (2011). Cogmed in your area. Retrieved from <http://www.cogmed.com/category/footer/cogmed-in-your-area>
- Cogmed. (2012). Cogmed working memory training. Retrieved from <http://www.cogmed.com>
- Colflesh, G. J. H., & Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review, 14*, 699–703. doi:10.3758/BF03196824
- Colom, R., Abad, F. J., Quiroga, M. Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence, 36*, 584–606. doi:10.1016/j.intell.2008.01.002
- Colom, R., Martínez-Molina, A., Chun Shih, P., & Santacreu, J. (2010). Intelligence, working memory, and multitasking performance. *Intelligence, 38*, 543–551. doi:10.1016/j.intell.2010.08.002
- Colom, R., Quiroga, M. Á., Shih, P. C., Martínez, K., Burgaleta, M., Martínez-Molina, A., . . . Ramírez, I. (2010). Improvement in working memory is not related to increased intelligence scores. *Intelligence, 38*, 497–505. doi:10.1016/j.intell.2010.06.008
- Conners, C. K. (2001). *Conners Rating Scales—revised*. North Tonawanda, NY: Multihealth Systems.
- Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998). The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology, 26*, 257–268. doi:10.1023/A:1022602400621
- Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8*, 331–335. doi:10.3758/BF03196169
- Conway, A. R. A., Getz, S. J., Macnamara, B., & Engel de Abreu, P. M. J. (2010). Working memory and fluid intelligence: A multi-mechanism view. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1016/j.intell.2010.07.003
- Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology, 47*, 738–748. doi:10.1037/0022-3514.47.4.738
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–114. doi:10.1017/S0140525X01003922
- Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42–100. doi:10.1016/j.cogpsych.2004.12.001
- Craik, F. I. M., & Birtwhistle, J. (1971). Proactive inhibition in free recall. *Journal of Experimental Psychology, 91*, 120–123. doi:10.1037/h0031835
- Crawford, J. R., Obonsawin, M. C., & Allan, K. M. (1998). PASAT and components of WAIS-R performance: Convergent and discriminant validity. *Neuropsychological Rehabilitation, 8*, 255–272. doi:10.1080/713755575
- Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica, 50*, 291–323. doi:10.1016/0001-6918(82)90044-0
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper and Row.
- Dahlin, E., Nyberg, L., Bäckman, L., & Stigsdotter Neely, A. (2008). Plasticity of executive functioning in young and old adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23*, 720–730. doi:10.1037/a0014296
- Dahlin, E., Stigsdotter Neely, A., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320*, 1510–1512. doi:10.1126/science.1155466
- Dahlin, K. I. E. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing, 2011*, 24, 479–491. doi:10.1007/s11145-010-9238-y
- Daneman, M., & Carpenter, P. A. (1980). Individual difference in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*, 450–466. doi:10.1016/S0022-5371(80)90312-6
- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433. doi:10.3758/BF03214546
- DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborcht, M., Troseth, G. L., . . . O'Doherty, K. (2010). Do babies learn from baby media? *Psychological Science, 21*, 1570–1574. doi:10.1177/0956797610384145
- Desoete, A. (2009). Mathematics and metacognition in adolescents and adults with learning disabilities. *International Electronic Journal of Elementary Education, 2*, 82–100. Retrieved from <http://www.iejee.com/index.html>
- de Vos, T. (1992). *Tempo test rekenen* [Tempo Test Arithmetic]. Lisse, the Netherlands: Swets Test Publishers.
- Diamond, A. (2005). Attention-deficit disorder (attention-deficit/hyperactivity disorder without hyperactivity): A neurobiologically and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Development and Psychopathology, 17*, 807–825. doi:10.1017/S0954579405050388
- Engel de Abreu, P. M. J., Conway, A. R. A., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence, 38*, 552–561. doi:10.1016/j.intell.2010.07.003
- Engel de Abreu, P. M. J., Gathercole, S. E., & Martin, R. (2011). Disentangling the relationship between working memory and language: The

- roles of short-term storage and cognitive control. *Learning and Individual Differences*, 21, 569–574. doi:10.1016/j.lindif.2011.06.002
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11, 19–23. doi:10.1111/1467-8721.00160
- Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in the role of working memory in comprehension and following directions. *Journal of Educational Research*, 84, 253–262.
- Engle, R. W., & Oransky, N. (1999). Multi-store versus dynamic models of temporary storage in memory. In R. Sternberg (Ed.), *The nature of cognition* (pp. 514–555). Cambridge, MA: MIT Press.
- 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
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143–149. doi:10.3758/BF03203267
- Fabbro, F., Muzur, A., Bellen, R., Calacione, R., & Bava, A. (1999). Effects of praying and a working memory task in participants trained in meditation and controls on the occurrence of spontaneous thoughts. *Perceptual and Motor Skills*, 88, 765–770. doi:10.2466/PMS.88.3.765-770
- Festinger, L. (1957). *Theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- French, J. R. P. (1953). Experiments in field settings. In L. Festinger & D. Katz (Eds.), *Research methods in the behavioral sciences* (pp. 98–135). New York, NY: Holt, Rinehart, and Winston.
- Gathercole, S. E., & Alloway, T. P. (2004). Working memory and classroom learning. *Dyslexia Review*, 15, 4–10.
- Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory deficits in laboratory analogues of activities. *Applied Cognitive Psychology*, 22, 1019–1037. doi:10.1002/acp.1407
- Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in 6- and 7-year-old children. *Journal of Educational Psychology*, 92, 377–390. doi:10.1037/0022-0663.92.2.377
- George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, 49, 1083–1094. doi:10.1016/j.neuropsychologia.2011.02.001
- Gibson, B. S., Gondoli, D. M., Flies, A. C., Dobrzanski, B. A., & Unsworth, N. (2009). Application of the dual-component model of working memory to ADHD. *Child Neuropsychology*, 16, 60–79. doi:10.1080/09297040903146958
- Gibson, B. S., Gondoli, D. M., Johnson, A. C., Steeger, C. M., Dobrzanski, 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*. doi:10.1080/09297049.2010.551186
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, 6, 316–322. doi:10.1038/nm1014
- Greenwald, A. G., Spangenberg, E. R., Pratkanis, A. R., & Eskenazi, J. (1991). Double-blind tests of subliminal self-help audiotapes. *Psychological Science*, 2, 119–122. doi:10.1111/j.1467-9280.1991.tb00112.x
- Gronwall, D. M. A. (1977). Paced auditory serial-addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367–373. doi:10.2466/pms.1977.44.2.367
- Güss, C. D., & Dörner, D. (2011). Cultural differences in dynamic decision-making strategies in a non-linear time-delayed task. *Cognitive Systems Research*, 12, 365–376. doi:10.1016/j.cogsys.2010.12.003
- Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, 18, 1279–1296. doi:10.1016/0042-6989(78)90218-3
- Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, 24, 1149–1167. doi:10.1002/acp.1624
- Healey, M. K., & Miyake, A. (2009). The role of attention during retrieval in working-memory span: A dual-task study. *The Quarterly Journal of Experimental Psychology*, 62, 733–745. doi:10.1080/17470210802229005
- Heitz, R. P., Redick, T. S., Hambrick, D. Z., Kane, M. J., Conway, A. R. A., & Engle, R. W. (2006). Working memory, executive function, and general fluid intelligence are not the same. *Behavioral and Brain Sciences*, 29, 135–136. doi:10.1017/S0140525X06319036
- 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. doi:10.1111/j.1467-7687.2009.00848.x
- 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. doi:10.1002/acp.1589
- Hossiep, R., Turck, D., & Hasella, M. (1999). *Bochumer matrizentest: BOMAT* [Bochum matrices–advanced–short version]. Göttingen, Germany: Hogrefe.
- Houben, K., Wiers, R. W., & Jansen, A. (2011). Getting a grip on drinking behavior: Training working memory to reduce alcohol abuse. *Psychological Science*, 22, 968–975. doi:10.1177/0956797611412392
- Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 645–662. doi:10.1037/0278-7393.33.4.645
- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 851–860. doi:10.1037/a0023437
- Ilkowska, M. (2011). *The relationship between higher-order cognition and personality* (Unpublished doctoral dissertation). Georgia Institute of Technology, Atlanta, GA.
- Jaeggi, S. M., Buschkuhl, M., Jonidas, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the USA*, 105, 6829–6833. doi:10.1073/pnas.0801268105
- Jaeggi, S. M., Buschkuhl, M., Jonidas, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the USA*, 108, 10081–10086. doi:10.1073/pnas.1103228108
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, 18, 394–412. doi:10.1080/09658211003702171
- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning: Implications for training and transfer. *Intelligence*, 38, 625–635. doi:10.1016/j.intell.2010.09.001
- Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-test*, [Berlin Intelligence Structure Test]. Göttingen, Germany: Hogrefe.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jones, J. D. (2007). The effects of music training and selective attention on working memory during bimodal processing of auditory and visual stimuli. *Dissertation Abstracts International: Section A. Humanities and Social Sciences*, 67, 2805.
- Jungle Memory. (2010). Who can benefit? Retrieved from [http://www.junglememory.com/pages/who\\_can\\_benefit](http://www.junglememory.com/pages/who_can_benefit)
- Kane, M. J., Bleckley, K. M., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of*

- Experimental Psychology: General*, 130, 169–183. doi:10.1037/0096-3445.130.2.169
- Kane, M. J., Brown, L. E., Little, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapił, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science*, 18, 614–621. doi:10.1111/j.1467-9280.2007.01948.x
- Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–46). New York, NY: Oxford University Press.
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the *N*-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 615–622. doi:10.1037/0278-7393.33.3.615
- Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47–70. doi:10.1037/0096-3445.132.1.47
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 66–71. doi:10.1037/0033-2909.131.1.66
- Kane, M. J., Hambrick, D. Z., Tuholski, 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
- Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of “executive attention.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 749–777. doi:10.1037/0278-7393.32.4.749
- Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science*, 12, 978–990. doi:10.1111/j.1467-7687.2009.00846.x
- Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3rd ed.). New York, NY: Wiley.
- Keulers, E. H. H., Hendriksen, J. G. M., Feron, F. J. M., Wassenberg, R., Wuisman-Frerker, M. G. F., Jolles, J., & Vlese, J. S. H. (2007). Methylphenidate improves reading performance in children with attention deficit hyperactivity disorder and comorbid dyslexia: An unblinded clinical trial. *European Journal of Paediatric Neurology*, 11, 21–28. doi:10.1016/j.ejpn.2006.10.002
- Kim, J., & Mueller, C. W. (1978). *Introduction to factor analysis: What it is and how to do it*. Beverly Hills, CA: Sage.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55, 352–358. doi:10.1037/h0043688
- Kleider, H. M., Parrott, D. J., & King, T. Z. (2010). Shooting behaviour: How working memory and negative emotionality influence police officer shoot decisions. *Applied Cognitive Psychology*, 24, 707–711. doi:10.1002/acp.1580
- Klingberg, T. (2009). *The overflowing brain: Information overload and the limits of working memory*. New York, NY: Oxford Press.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, 14, 317–324. doi:10.1016/j.tics.2010.05.002
- Klingberg, T., Fernell, E., Olesen, P., 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 & Adolescent Psychiatry*, 44, 177–186. doi:10.1097/00004583-200502000-00010
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, 24, 781–791. doi:10.1076/jcen.24.6.781.8395
- Kronenberger, W. G., Pisoni, D. B., Henning, S. C., Colson, B. G., & Hazzard, L. M. (2011). Working memory training improves memory capacity and sentence repetition skills in deaf children with cochlear implants: A pilot study. *Journal of Speech, Language, and Hearing Research*, 54, 1182–1196. doi:10.1044/1092-4388(2010/10-0119)
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity? *Intelligence*, 14, 389–433. doi:10.1016/S0160-2896(05)80012-1
- Lachman, M. E., Neupert, S. D., Bertrand, R., & Jette, A. M. (2006). The effects of strength training on memory in older adults. *Journal of Aging and Physical Activity*, 14, 59–73.
- Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction*, 17, 336–344. doi:10.1016/j.learninstruc.2007.02.010
- Lemmink, K. A., & Visscher, C. (2005). Effect of intermittent exercise on multiple-choice reaction times of soccer players. *Perceptual and Motor Skills*, 100, 85–95. doi:10.2466/pms.100.1.85-95
- Li, S.-C., Schmiedeck, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, 23, 731–742. doi:10.1037/a0014343
- Lichtenberger, E. O. (2005). General measures of cognition for the pre-school child. *Mental Retardation and Developmental Disabilities Research Reviews*, 11, 197–208. doi:10.1002/mrdd.20076
- Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path and structural equation analysis* (4th ed.). Hillsdale, NJ: Erlbaum.
- Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, 7, 166–174. doi:10.3758/BF03197535
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, 136, 659–676. doi:10.1037/a0020080
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203. doi:10.1037/0033-2909.109.2.163
- MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, 4, 383–391. doi:10.1016/S1364-6613(00)01530-8
- Maguire, E. A., Valentine, E. R., Wilding, J. M., & Kapur, N. (2003). Routes to remembering: The brains behind superior memory. *Nature Neuroscience*, 6, 90–95. doi:10.1038/nn988
- Mayo, E. (1933). *The human problems of an industrial civilization*. New York, NY: Macmillan.
- McArdle, J. J., & Prindle, J. J. (2008). A latent change score analysis of a randomized clinical trial in reasoning training. *Psychology and Aging*, 23, 702–719. doi:10.1037/a0014349
- McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, 24, 222–243. doi:10.1037/a0017619
- McCarney, R., Warner, J., Illife, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne effect: A randomized, controlled trial. *BMC Medical Research Methodology*, 7(30). doi:10.1186/1471-2288-7-30
- McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, 11, 103–107. doi:10.1038/nn2024
- McNab, F., Varrone, A., Farde, L., Jucaite, A., Bystritsky, P., Forssberg, H., & Klingberg, T. (2009). Changes in cortical dopamine D1 receptor

- binding associated with cognitive training. *Science*, *323*, 800–802. doi:10.1126/science.1166102
- McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, *29*, 10–17. doi:10.3758/BF03195736
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory Capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 196–204. doi:10.1037/a0014104
- Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science*, *21*, 914–919. doi:10.1177/0956797610373933
- Mezzacappa, E., & Buckner, J. C. (2010). Working memory training for children with attention problems or hyperactivity: A school-based pilot study. *School Mental Health*, *2*, 202–208. doi:10.1007/s12310-010-9030-9
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97. doi:10.1037/h0043158
- Mindspark. (2011). *Brain fitness pro*. Retrieved from <http://www.mindspark.com>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. doi:10.1006/cogp.1999.0734
- 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. doi:10.1037/0096-3445.130.4.621
- Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence*, *37*, 327–328. doi:10.1016/j.intell.2009.04.005
- Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 703–713. doi:10.1037/0278-7393.31.4.703
- 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. doi:10.3758/s13423-010-0034-0
- Nigg, J. T. (2010). Attention-deficit/hyperactivity disorder: Endophenotypes, structure, and etiological pathways. *Current Directions in Psychological Science*, *19*, 24–29. doi:10.1177/0963721409359282
- Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence—their correlation and their relation: A comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 61–65. doi:10.1037/0033-2909.131.1.61
- Oken, B. S., Flegal, K., Zajdel, D., Kishiyama, S., Haas, M., & Peters, D. (2008). Expectancy effect: Impact of pill administration on cognitive performance in healthy seniors. *Journal of Clinical and Experimental Neuropsychology*, *30*, 7–17. doi:10.1080/13803390701775428
- Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*, 75–79. doi:10.1038/nn1165
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776–783. doi:10.1037/h0043424
- Orne, M. T. (1972). Communication by the total experimental situation: Why it is important, how it is evaluated, and its significance for the ecological validity of findings. In P. Pliner, L. Krames, & T. Alloway (Eds.), *Communication and affect: Language and thought* (pp. 157–191). New York, NY: Academic Press.
- Orne, M. T., & Scheibe, K. E. (1964). The contribution of nondeprivation factors in the production of sensory deprivation effects: The psychology of the “panic button.” *Journal of Abnormal and Social Psychology*, *68*, 3–12. doi:10.1037/h0048803
- Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Fricke, D., Zwahr, M., & Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging*, *11*, 621–637. doi:10.1037/0882-7974.11.4.621
- Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–183). Hillsdale, NJ: Erlbaum.
- Persson, J., & Reuter-Lorenz, P. A. (2008). Gaining control: Training of executive function and far transfer of the ability to resolve interference. *Psychological Science*, *19*, 881–888. doi:10.1111/j.1467-9280.2008.02172.x
- Persson, J., & Reuter-Lorenz, P. A. (2011). Retraction of “Gaining control: Training of executive function and far transfer of the ability to resolve interference.” *Psychological Science*, *22*, 562. doi:10.1177/0956797611404902
- Pollack, I., Johnson, L. B., & Knaff, P. R. (1959). Running memory span. *Journal of Experimental Psychology*, *57*, 137–146. doi:10.1037/h0046137
- Raven, J. C. (1990). *Advanced progressive matrices*. Oxford, United Kingdom: Oxford Psychological Press.
- Raven, J. C. (1995). *Coloured progressive matrices*. Oxford, United Kingdom: Oxford Psychologists Press.
- Raven, J. C. (1998). *Raven standard progressive matrices plus*. Oxford, United Kingdom: Oxford Psychologists Press.
- Richmond, L. L., Morrison, A. B., Chein, J. M., & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and Aging*, *26*, 813–822. doi:10.1037/a0023631
- Robertson, I., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The test of everyday attention*. Bury St. Edmunds, England: Thames Valley Test.
- Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, *3*, 176–179. doi:10.1111/1467-8721.ep10770698
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York, NY: Hold, Rinehart, and Winston.
- Rosnow, R. L., & Rosenthal, R. (1997). *People studying people: Artifacts and ethics in behavioral research*. New York, NY: Freeman.
- Rueda, M. R., Rothbard, M. K., McCandliss, B. D., Saccamanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences, USA*, *102*, 14931–14936. doi:10.1073/pnas.0506897102
- Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age difference in working memory. *Developmental Psychology*, *27*, 763–776. doi:10.1037/0012-1649.27.5.763
- Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review*, *15*, 364–371. doi:10.3758/PBR.15.2.364
- Schmeichel, B. J., Volokhov, R., & Demaree, H. A. (2008). Working memory capacity and the self-regulation of emotional expression and experience. *Journal of Personality and Social Psychology*, *95*, 1526–1540. doi:10.1037/a0013345
- Schmiedek, F., Hildebrandt, A., Lövdén, M., Lindenberger, U., & Wilhelm, O. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1089–1096. doi:10.1037/a0015730



- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience, 2*, 1–10. doi:10.3389/fnagi.2010.00027
- Seidler, R. D., Bernard, J. A., Buschkuhl, M., Jaeggi, S., Jonides, J., & Humfleet, J. (2010). *Cognitive training as an intervention to improve driving ability in the older adult* (Technical Report No. M-CASTL 2010–01). Ann Arbor: University of Michigan.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- 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
- Shavelson, R. J., Yuan, K., & Alonzo, A. (2008). On the impact of computer training on working memory and fluid intelligence. In D. C. Berliner & H. Kuerminz (Eds.), *Fostering change in institutions, environments, and people: A festschrift in honor of Gavriel Salomon* (pp. 35–48). New York, NY: Routledge.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50*, 245–276.
- Shute, V. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research, 7*, 1–24. doi:10.2190/VQJD-T1YD-5WVB-RYPJ
- Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic Bulletin & Review, 14*, 840–845. doi:10.3758/BF03194109
- Sohlberg, M. M., McLaughlin, K. A., Pavese, A., Heidrich, A., & Posner, M. I. (2000). Evaluation of attention process training and brain injury education in persons with acquired brain injury. *Journal of Clinical and Experimental Neuropsychology, 22*, 656–676. doi:10.1076/1380-3395(200010)22:5;1-9;FT656
- St. Clair-Thompson, H. L. (2010). Backwards digit recall: A measure of short-term memory or working memory? *European Journal of Cognitive Psychology, 22*, 286–296. doi:10.1080/09541440902771299
- St. Clair-Thompson, H. L., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology, 30*, 203–219. doi:10.1080/01443410903509259
- Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences, USA, 105*, 6791–6792. doi:10.1073/pnas.0803396105
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662. doi:10.1037/h0054651
- Süß, H. M., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence, 30*, 261–288. doi:10.1016/S0160-2896(01)00100-3
- Swanson, H. L. (2003). Age-related differences in learning disabled and skilled readers' working memory. *Journal of Experimental Child Psychology, 85*, 1–31. doi:10.1016/S0022-0965(03)00043-2
- Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for math disabilities. *Journal of Education Psychology, 96*, 471–491. doi:10.1037/0022-0663.96.3.471
- Tang, Y.-Y., & Posner, M. I. (2009). Attention training and attention state training. *Trends in Cognitive Sciences, 13*, 222–227. doi:10.1016/j.tics.2009.01.009
- te Nijenhuis, J., van Vianen, A. E. M., & van der Flier, H. (2007). Score gains on g-loaded tests: No g. *Intelligence, 35*, 283–300. doi:10.1016/j.intell.2006.07.006
- Thorell, L. B., Lindqvist, S., Bergman, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12*, 106–113. doi:10.1111/j.1467-7687.2008.00745.x
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in on mental function upon the efficiency of other functions (I). *Psychological Review, 8*, 247–261. doi:10.1037/h0074898
- Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1*, 86–98. doi:10.1016/0010-0285(70)90006-X
- Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49*, 446–468. doi:10.1016/S0749-596X(03)00095-0
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154. doi:10.1016/0749-596X(89)90040-5
- 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). Individual differences in working memory capacity and retrieval: A cue-dependent search approach. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 241–258). New York, NY: Psychology Press.
- 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 ability. *Psychological Bulletin, 133*, 1038–1066. doi:10.1037/0033-2909.133.6.1038
- Unsworth, N., & Engle, R. W. (2007c). 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., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1302–1321. doi:10.1037/0278-7393.30.6.1302
- Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention, memory, or both? A direct test of the dual-component model. *Journal of Memory and Language, 62*, 392–406. doi:10.1016/j.jml.2010.02.001
- Unsworth, N., Spillers, G. J., & Brewer, G. A. (2009). Examining the relations among working memory capacity, attention control, and fluid intelligence from a dual-component framework. *Psychology Science Quarterly, 4*, 388–402.
- Van der Molen, M. J., Van Luit, J. E. H., Van der Molen, M. W., Klugkist, I., & Jongmans, M. J. (2010). Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research, 54*, 433–447. doi:10.1111/j.1365-2788.2010.01285.x
- van Vugt, M. K., & Jha, A. P. (2011). Investigating the impact of mindfulness meditation training on working memory: A mathematical modeling approach. *Cognitive, Affective, & Behavioral Neuroscience, 11*, 344–353. doi:10.3758/s13415-011-0048-8
- Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., & Gruzelier, J. (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *International Journal of Psychophysiology, 47*, 75–85. doi:10.1016/S0167-8760(02)00091-0
- Vogt, A., Kappos, L., Calabrese, P., Stöcklin, M., Gschwind, L., Opwis, K., & Penner, I. (2009). Working memory training in patients with multiple sclerosis: Comparison of two different training schedules. *Restorative Neurology and Neuroscience, 27*, 225–235. doi:10.3233/RNN-2009-0473
- Wechsler, D. (1993). *Wechsler Objective Reading Dimensions*. New York, NY: Psychological Corporation.

- Wechsler, D. (1995). *Wechsler Preschool and Primary Scale of Intelligence—Revised*. New York, NY: Psychological Corporation.
- Wechsler, D. (1996). *Wechsler Objective Number Dimensions*. New York, NY: Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. London, England: Harcourt Assessment.
- Wechsler, D. (1999). *Manual for the Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2002). *WPPSI-III technical and interpretation manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2004). *Wechsler Preschool and Primary Scale of Intelligence—(3rd ed.)*. New York, NY: Psychological Corporation.
- Westerberg, H., Brehmer, Y., D'Hondt, N., Söderman, D., & Bäckman, L. (2008). *Computerized training of working memory: A controlled, randomized trial*. Poster presented at the meeting of the Cognitive Neuroscience Society, San Francisco, CA.
- Westerberg, H., Hirvikoski, T., Forsberg, H., & Klingberg, T. (2004). Visuo-spatial working memory span: A sensitive measure of cognitive deficits in children with ADHD. *Child Neuropsychology, 10*, 155–161. doi:10.1080/09297040490911014
- Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Östenson, M.-L., Bartfai, A., & Klingberg, T. (2007). Computerized working memory training after stroke: A pilot study. *Brain Injury, 21*, 21–29. doi:10.1080/02699050601148726
- Wexler, B. E., Anderson, M., Fulbright, R. K., & Gore, J. C. (2000). Preliminary evidence of improved verbal working memory performance and normalization of task-related frontal lobe activation in schizophrenia following cognitive exercises. *American Journal of Psychiatry, 157*, 1694–1697. doi:10.1176/appi.ajp.157.10.1694
- Wiseman, R., & Schlitz, M. (1997). Experimenter effects and the remote detection of staring. *Journal of Parapsychology, 61*, 197–208.
- Yntema, D. B. (1963). Keeping track of several things at once. *Human Factors, 5*, 7–17. doi:10.1016/0001-6918(67)90075-3
- Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: Evidence of brief mental training. *Consciousness and Cognition, 19*, 597–605. doi:10.1016/j.concog.2010.03.014
- Zhao, X., Wang, Y. X., Liu, D. W., & Zhou, R. L. (2011). Effect of updating training on fluid intelligence in children. *Chinese Science Bulletin, 56*, 2202–2205. doi:10.1007/s11434-011-4553-5

## Appendix

### Reliability Information for Transfer Tasks

Test	Reliability	Source	Description
ADHD			
Conners Rating Scale	Cronbach's $\alpha = .75-.94$	Conners et al. (1998)	Rating inventory
Attention controlled/sustained			
Choice Reaction Time	Cronbach's $\alpha = .99$	Lemmink & Visscher (2005)	Information-dependent response
Continuous Performance Task	Test-retest = $.73-.86$	Borgaro et al. (2003)	Respond when specific item presented
Go/No-Go Task	Cronbach's $\alpha = .95$	McVay & Kane (2009)	Respond or withhold response, depending upon item
Paced Auditory Serial-Attention Task (PASAT)	Cronbach's $\alpha = .90$	Crawford, Obonsawin, & Allen (1998)	Serial digits, add current to previous digit
Stroop Task (no congruent trials)			
Response time	Split-half = $.55$	Hutchison (2007)	State hue in which an incongruent word is printed
Errors	Split-half = $.48$	Hutchison (2007)	State hue in which an incongruent word is printed
Test of Everyday Attention (TEA)	Test-retest = $.75-.78$	Clayton et al. (1999)	Distracted performance and multitasking
Fluid intelligence/achievement			
Berlin Intelligence Structure Test (BIS)			
Figural	Cronbach's $\alpha = .88$	Jäger, Süß, & Beauducel (as cited in Süß et al., 2002)	Visual analogies, rule discovery, font recognition
Numerical	Cronbach's $\alpha = .90$	Jäger, Süß, & Beauducel (as cited in Süß et al., 2002)	Sequence completion, arithmetic, categorization
Verbal	Cronbach's $\alpha = .90$	Jäger, Süß, & Beauducel (as cited in Süß et al., 2002)	Reasoning with verbal materials

(Appendix continues)

## Appendix 1 (continued)

Test	Reliability	Source	Description
Bochumer Matrizen-Test (BOMAT)	Cronbach's $\alpha = .58$	Jaeggi, Studer-Luethi, et al. (2010)	Select missing component of a grid-based series
Differential Aptitude Test (DAT)	Cronbach's $\alpha = .79-.86$	Colom et al. (2008)	Select missing component of a grid-based series
Abstract Reasoning			
Spatial Relations	Cronbach's $\alpha = .86$	Colom et al. (2008)	Mental object folding
Verbal Reasoning	Cronbach's $\alpha = .80-.86$	Colom et al. (2008)	Reasoning with verbal materials
Eén Minuut Test	Test-retest = .91	Keulers et al. (2007)	Word reading
Raven's Progressive Matrices	Cronbach's $\alpha = .76$	Broadway & Engle (2010)	Select missing component of a grid-based series
Tempo Test Rekenen	Cronbach's $\alpha = .90$	Desoete (2009)	Arithmetic operations
Test of Nonverbal Intelligence (TONI)	Cronbach's $\alpha = .85$	Güss & Dörner (2011)	Select missing component of a grid-based series
Wechsler Abbreviated Scales of Intelligence (WASI)			
Performance IQ	Split-half = .94-.96	Wechsler (as cited in Kaufman & Lichtenberger, 2006)	Select missing component of a grid-based series
Verbal IQ	Split-half = .93-.96	Wechsler (as cited in Kaufman & Lichtenberger, 2006)	Vocabulary and word knowledge
Wechsler Preschool and Primary Scale of Intelligence (WPPSI)			
Block Design	Internal consistency = .75	Wechsler (as cited in Lichtenberger, 2005)	Arrange blocks to reproduce specific pattern
Wechsler Objective Number Dimensions (WOND)			
Mathematical Reasoning	Test-retest = .89	Alloway et al. (2009)	Geometric shape identification and word problems
Numerical Operations	Test-retest = .85	Alloway et al. (2009)	Number identification and computation
Wechsler Objective Reading Dimensions (WORD)			
Basic Reading	Test-retest = .95	Alloway et al. (2009)	Letter naming, word reading
Reading Comprehension	Test-retest = .92	Alloway et al. (2009)	Word matching, sentence reading with questions
Spelling	Test-retest = .91	Alloway et al. (2009)	Writing words and sounds
Working Memory Capacity/Short-Term Memory			
Complex span			
Operation span	Cronbach's $\alpha = .80$	Kane et al. (2004)	Remember lists of letters, while solving mathematical equations
Symmetry span	Cronbach's $\alpha = .86$	Kane et al. (2004)	Remember spatial locations, while making symmetry judgments
<i>n</i> -back			
Single <i>n</i> -back	Cronbach's $\alpha = .78$	Jaeggi, Studer-Leuthi, et al. (2010)	Respond when an item matches an item presented <i>n</i> ago
Dual <i>n</i> -back	Cronbach's $\alpha = .91$	Jaeggi, Studer-Leuthi, et al. (2010)	Respond when one of two items matches an item presented <i>n</i> ago
Running Memory Span	Cronbach's $\alpha = .76-.86$	Broadway & Engle (2010)	Remember the final <i>n</i> items of a list
Simple Span			
Forward (nonrhyming words)	Cronbach's $\alpha = .74$	Engle et al. (1999)	Remember a short list of words
Backward (nonrhyming words)	Cronbach's $\alpha = .69$	Engle et al. (1999)	Remember a short list of words in reverse
Corsi Block (visuo-spatial)	Cronbach's $\alpha = .83$	Colom et al. (2008)	Remember a series of spatial locations

Note. Split-half = split-half reliability; Test-retest = Test-retest reliability.

Received February 15, 2011  
Revision received November 25, 2011  
Accepted December 12, 2011 ■