A Randomized-Controlled Trial of Working Memory Training in Youth with ADHD

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Christine Adelaide Hanson

Graduate Program in Psychology

The Ohio State University

2013

Dissertation Committee:

Steven Beck, Advisor

Keith Yeates

Ruchika Prakash

Abstract

The present study investigated whether working memory training (WMT) would improve working memory (WM), planning/organization, executive functioning, attention, hyperactivity/impulsivity, and reading comprehension in individuals with Attention Deficit Hyperactivity Disorder (ADHD). Twenty-eight children and adolescents with ADHD completed WMT, which consisted of 25 sessions lasting 30-45 minutes completed over about 6 weeks. Participants were randomly assigned to either a difficult adaptive WMT program or a control program, which maintained a low-level of difficulty. We predicted that the experimental group would show greater improvements than the active control group. The experimental group showed a trend towards improving more than the active control group on nonverbal short-term memory (STM), one measure of verbal WM, parent-rated inattention, After WMT participants in both groups improved on verbal STM, nonverbal STM, nonverbal WM, one measure of verbal WM, parent-reported WM, a WM composite, parent-rated inattention, reading comprehension, one participant-administered measure of planning/organization, parent-rated planning/organization, and parent-rated executive functioning. Participants did not improve on one measure of verbal WM, parent-rated hyperactivity/impulsivity, and a participant-administered measure of attention, one participant-administered measure of planning/organization, and a participant-administered measure of executive functioning.

There was not enough teacher-report data to come to any meaningful conclusions. This lends some support that WMT can lead to improvements in broad cognitive functions; however, pre-treatment to post-treatment improvements may have been due to practice or expectancy effects. It is unclear whether the training needs to be difficult and adaptive in order to lead to improvements or if just training WM for a certain period of time is sufficient to lead to benefits. Future studies need to investigate the necessary components of WMT and whether the improvements following WMT are clinically significant, stable over time, and not just due to practice effects, rater expectancy effects, or regression to the mean. Additional replication studies are needed showing improvements in cognitive and academic functions following WMT. Future studies should investigate whether certain WMT programs lead to improvements in certain cognitive and academic functions.

This is dedicated to my husband, my family, and my friends. Without your unending support I never would have believed I could have accomplished this. I thank you for sticking with me through it all. Through all the long days, sacrifices, tears, frustration, missed family functions, set-backs, triumphs, and distance between us you have always given me unconditional love and friendship, encouraging words, celebrations, ears to listen, laughter through the tears, a balanced perspective, reminders that it is worth it, and hope for the future. Thank you for enduring this marathon with me. We have finally crossed the finish line. I can finally come back to you and back to my home, Wisconsin. I cannot wait!

Acknowledgments

I would like to thank my advisor, Dr. Steve Beck, for beginning this line of research at Ohio State, helping with the study design, acquisition of measures and other resources needed for the study, recruitment of participants, invaluable comments on my dissertation, and, most importantly, for believing in me every step of the way. I would like to thank my labmate, Synthia Puffenberger, for being my partner in conducting this study and helping every step of the way. I would especially like to thank Synthia for her much needed organizational abilities, which were used for completing the arduous task of scheduling all of the data collection meetings. I would like to thank all of the study participants, their parents, their teachers, and other school personnel who allowed for data collection. I would like to thank the people at Cogmed for donating the working memory training programs and promptly replying to any and all tech support needs. I would like to thank Dr. Bill Benninger for being our liaison with Cogmed. Lastly, I would like to thank our research assistants, Megan French, Andrea Hatfield, Halle Stolarsky, Erica Reams, Michael Eastman, Brittney Gallt, Alyssa Price, and Akua Adu-gyamfi, for an excellent job in learning the standardized administration of measures, administering measures to participants and their parents, scoring the measures, and entering the data.

Vita

June 2003 .....Elkhorn Area High School

May 2007 .....B.S. Psychology, University of Wisconsin Madison

May 2007 .....B.S. Spanish, University of Wisconsin Madison

Sept 2007 .....Social and Behavioral Sciences Fellowship, The Ohio State University

Oct 2007-Sept 2008 .....University Fellowship, The Ohio State University

Sep 2008-Dec 2008.....Graduate Teaching Associate, Department of Psychology, The Ohio State University

Dec 2008-June 2009 .....Clinic Coordinator**,** OSU Psychological Services Center, The Ohio State University

July 2009-Sept 2009 .....Social and Behavioral Sciences Fellowship, The Ohio State University

Sept 2009-June 2012.....Graduate Teaching Associate, Department of Psychology, The Ohio State University

Dec 2009 .....M.A. Clinical Psychology, The Ohio State University

June 2012 to present.....Pre-doctoral Intern, Psychiatry Department, Geisinger Medical Center

Publications

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

Fields of Study

Major Field: Psychology

Table of Contents

List of Tables

Chapter 1: Introduction

Working Memory

Working memory (WM) is a system that allows one to temporarily hold information in mind long enough to use the information for some purpose (Baddeley, 2000). WM has limited capacity, in that only a limited amount of information can be maintained and processed at one time (Repovš & Baddeley, 2006). Some theorists argue that WM should be thought of as distinct from short-term memory (STM) in that it involves the manipulation of the information held in mind or its maintenance in the face of interference, rather than just the passive storage of this information (Unsworth & Engle, 2007; Baddeley & Hitch, 1994; Engle, Tuholski, Laughlin & Conway, 1999). There is evidence to support that WM and STM are distinct. For example, when six STM and six WM tasks were administered to a group of children, a factor analysis found that the best fitting model included separate, although correlated, factors for STM and WM (Alloway, Gathercole & Pickering, 2006). Two latent variable studies with adults also found separate factors for WM and STM, although these factors were moderately correlated (Engle et al., 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne & Engle, 2004). Research has also found that brain activation is somewhat different for tasks that involve only storage versus tasks that require storage and manipulation, in that

manipulation involves more dorsolateral prefrontal cortex activation (D'Esposito, Postle, Ballard & Lease, 1999; Wager & Smith, 2003). Therefore, the research appears to support that WM and STM are distinct, although related, constructs.

The model of WM originally proposed by Baddeley and Hitch (1974) is often used as a theoretical framework for research on WM. The revised version of this model proposes that WM is broken down into four components: the phonological loop, the visuo-spatial sketchpad, the central executive, and the episodic buffer (Repovš & Baddeley, 2006). The phonological loop is conceptualized as the part of WM that is responsible for holding verbal information in mind and primarily uses rehearsal (repeating the information over and over to oneself) as the means for remembering. The visuo-spatial sketchpad is thought to be responsible for holding non-verbal information in mind (Repovš & Baddeley, 2006). The central executive is described as regulating the system in that it directs attention, guides the flow of information, coordinates the execution of two or more tasks at once, and interacts with long-term memory (Repovš & Baddeley, 2006). Therefore, the central executive is likely not involved in tasks that require only storage of information (STM tasks), but is involved when the information is to be manipulated in some way or maintained in the face of interference (WM tasks). Lastly, the episodic buffer is able to store integrated verbal and visual-spatial information from WM and long-term memory (Repovš & Baddeley, 2006).

Activation of the prefrontal cortex, parietal regions, and the basal ganglia are important for WM tasks (Wager & Smith, 2003). Neuroimaging studies, using positron emission tomography (PET), indicate that the different components of WM may be

associated with different brain regions (Smith & Jonides, 1997). More complex WM tasks, which involve the central executive, rely more heavily on the frontal lobes, in particular the dorsolateral prefrontal cortex (DLPFC; Wager & Smith, 2003; Smith & Jonides, 1997). Dopamine (DA) functioning, in particular functioning of the D1 receptor, has also been implicated in WM. Studies have found that dopaminergic activity, specifically DA binding to D1 receptors, is important for WM tasks (Bäckman, Karlsson, Fischer, Karlsson, Brehmer, Rieckmann et al., 2011). Decreased binding of D1 receptors in the dorsolateral prefrontal cortex when completing WM tasks has been associated with worse performance on the WM tasks (Bäckman et al., 2011).

Normally, WM develops gradually from early childhood through adolescence. The average WM span of a preschooler is about one-third of the WM span of a teenager or young adult (Dempster, 1981). The average four-year-old has a verbal WM span of two to three items, whereas an average teenager's verbal WM span is around seven items (Gathercole & Baddeley, 1993). A similar two-to-three-fold increase in visuo-spatial WM span is also seen from the preschool to teenage years. Improvements in the functioning of the central executive are also seen throughout childhood (Gathercole & Baddeley, 1993). Gathercole and Baddeley (1993) suggest that this increased capacity is due to the increased efficacy of the WM system, rather than a qualitative change in the WM system.

One study compared performance on complex WM tasks in younger children (ages 8-12), adolescents (ages 13-17) and young adults (ages 18-25) (Crone, Wendelken, Donohue, van Leijenhorst, & Bunge, 2006). Brain activity was recorded with functional

magnetic resonance imaging (fMRI) while participants engaged in a WM task. This study found that younger children did worse than adolescents and adults on WM task and that younger child, in contrast to the adolescents and adults, did not show as much activation in the dorsolateral prefrontal cortex (DLPFC; Crone et al., 2006). Activation of the DLPFC was also positively correlated with performance on the WM tasks, such that those with more DLPFC activity do better on WM tasks (Crone et al., 2006). Other studies have found that developmental improvements are positively correlated with brain activity in the prefrontal cortex and parietal cortices (Klingberg, Forssberg & Westerberg, 2002a). This may indicate that decreased activation in frontal and parietal cortices is associated with poorer performance on WM tasks.

WM Effects Other Abilities

WM abilities, especially functioning of the central executive, are important for many areas of functioning, such as learning, reasoning, intelligence, and other cognitive functions (Baddeley, 2003; Kyllonen & Christal, 1990; Buehner, Krumm & Pick, 2005). Research suggests that WM abilities may be important for attention in that WM may be necessary to maintain the prioritization of relevant information or to keep in mind which stimuli are relevant (Desimone & Duncan, 1995; de Fockert, Rees, Frith & Lavie, 2001). This may then aid in directing attention to relevant rather than irrelevant stimuli. Therefore, this maintenance in WM of which stimuli are important may decrease the influence of distracters and, thus, decrease distractibility. Further evidence for the relationship between WM and attention comes from the fact that WM performance in

children with ADHD significantly correlates with symptoms of inattention (Martinussen & Tannock, 2006). Martinussen & Tannock (2006) found that inattentive symptoms were a significant predictor of verbal and visual spatial WM performance. In contrast, hyperactive symptoms did not significantly predict WM performance (Martinussen & Tannock, 2006). There is research to suggest that WM processes and attention also rely on the same frontoparietal neural network (Klingberg, 2010). Therefore neurological changes in these neural networks as a result of WM Training (WMT) would be predicted to lead to behavioral changes in other tasks that also rely on these neural networks, such as attention.

Recent evidence has suggested that WM is important for reasoning/fluid intelligence. WM tasks are often included on measures of intelligence and thought to contribute to overall intelligence, in particular fluid intelligence. In fact, studies have found that WM capacity accounts for about 36% of the variance in reasoning/fluid intelligence (Halford, Cowan & Andrews, 2007). Structural equation modeling has also found that WM has a strong effect on fluid reasoning and accounts for about 40% of the variance in fluid reasoning (Kane et al., 2004). This leads to the hypothesis that by improving WM, fluid intelligence may also be affected.

Planning is also considered to involve WM (Nigg, 2006). Performance on WM tasks significantly correlates with performance on planning tasks (St Clair-Thompson, 2011). Working memory ability, specifically the central executive domain of WM, significantly predicts planning ability (Badcock, Michie & Rock, 2005). When adults were required to complete the Tower of London, a common neuropsychological measure

of planning/organization, while completing either a verbal or visuo-spatial WM task involving the central executive, they did worse on the Tower of London (Phillips, Wynn, Gloomy, Sala & Logie, 1999). Therefore, it appears that the ability to maintain and manipulate information in memory is important for success on planning tasks. Activation of the DLPFC is seen in tasks involving planning as well as tasks involving WM and when the activation during these two tasks are compared, there are no significant differences in activation seen in the DLPFC (Owen, Doyon, Petrides & Evans, 1996). Given that similar brain regions are involved in both WM and planning tasks, improvements in WM after training, may also lead to improvements in planning/organization. Beck, Hanson, Puffenberger, Benninger, & Benninger (2010) found that after WMT, participants significantly improved on parent reported planning/organization. However, no studies to date have investigated whether WMT leads to improvements on neuropsychological measures of planning/organization.

Given that WM is implicated in many tasks, it follows that maintaining information in WM may be necessary for most executive functioning tasks and that this common need for WM may account for the correlations among different executive functions (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000). This implies that WM ability is essential for many tasks and that WM ability could potentially affect many other abilities. Therefore, the possibility of improving WM could have broad implications for improvements in many other areas of executive functioning.

WM ability has also been shown to be important for success on a variety of school related tasks (de Jong, 1998; Gathercole & Pickering, 2000; Swanson & Sachse-

Lee, 2001). WM predicts reading, spelling, and mathematics abilities (Alloway & Alloway, 2010). Alloway and Alloway (2010) found that WM abilities at five years of age predicted academic skills six years later at age eleven and that WM explained additional variance in these skills above and beyond intelligence. When looking at children whose WM abilities are in the bottom 10th percentile, 68-75% of them score in the borderline or impaired range on assessments of reading (Gathercole & Alloway, 2008). Those with reading disabilities are thought to have difficulties with both the phonological loop and central executive components of WM (de Jong, 2006). In recent meta-analyses, those with reading disabilities and ADHD scored significantly worse than those without on WM tasks (Martinussen et al., 2005; Hanson, 2011). In particular, WM has been shown to be important for reading comprehension (Cain, Oakhill, & Bryant, 2004; Swanson, Howard, & Sáez, 2006). Verbal WM ability significantly correlates with reading comprehension and the size of the correlation is similar to what is seen for the correlation between reading comprehension and other facets of reading, such as vocabulary, verbal abilities, and reading fluency (Cain et al., 2000; Seigneuric, Ehrlich, Oakhill & Yuill, 2000; Vukovic & Siegal, 2005). Additionally, WM abilities explains unique variance in reading comprehension even after controlling for things that are known to affect comprehension such as age, intelligence, vocabulary, word recognition, phonological awareness, rapid naming, and reading fluency (Cain, Oakhill & Bryant, 2000; Seigneuric et al., 2000; Cane et al., 2004; Vukovic & Siegal, 2005). Because of the involvement of WM in reading comprehension, it is possible that improvements in WM will also lead to improvements in reading comprehension. There is some evidence for

this hypothesis from longitudinal studies, which have found that improvements in WM predict improvements in reading comprehension (Dufva, Niemi & Voeten, 2001; Vukovic & Siegal, 2005).

Attention Deficit Hyperactivity Disorder

In addition to the core symptoms of inattention and/or hyperactivity/impulsivity seen in individuals with ADHD, ADHD is also associated with many other features including frontal lobe abnormalities, deficits in executive functions such as WM, planning/organization, and inhibition, and increased risk for learning problems, in particular reading disorders (Barkley, 2006; Nigg, 2006; Willcutt & Pennington, 2000). Although executive functioning deficits have been suggested in ADHD, specific executive function deficits, such as WM, are not consistently found in all individuals with ADHD. Jacob and Lesch (2006) reported different individuals with ADHD may present with different executive functioning deficits. Therefore, these executive functioning deficits are not thought to be necessary or sufficient to explain each case of ADHD (Jacob & Lesch, 2006). Nonetheless, these executive functioning deficits, specifically WM, may be useful in understanding and treating ADHD.

Many studies support that WM is impaired in individuals with ADHD and recent theoretical explanations for ADHD include a WM component (Barkley, 1997; Castellanos & Tannock, 2002). Barkley (1997) proposes that behavioral inhibition, which involves the capability to inhibit the prepotent response, is the central deficit in children with ADHD. He proposes that behavioral inhibition provides the delay

necessary for executive functions, including WM, to occur. He states that deficits in behavioral inhibition can lead to secondary impairments in some types of neuropsychological functioning, one of which is WM. These secondary impairments, such as WM, can then lead to decreased control of behavior by internal representations of information and decreased self-directed actions. In turn, this decrease in the control of behavior by internal information may manifest itself in the form of symptoms of inattention (Barkley, 1997).

Various studies have found deficits in both verbal and visual WM in children (as young as preschoolers) and adults with ADHD (Martinussen, Hayden, Hogg-Johnson & Tannock, 2005; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005; Hanson, 2011; Mariani & Barkley, 1997). Recent meta-analyses have reviewed hundreds of studies comparing WM abilities of those with ADHD to non-clinical control groups (Martinussen et al., 2005; Willcutt et al., 2005; Hanson, 2011). These meta-analyses generally have found moderate effect sizes for the difference on WM abilities in those with versus those without ADHD. There is also evidence that all three components of Baddeley's model of WM: the phonological loop, visuo-spatial sketchpad, and the central executive, are impaired in ADHD (Rapport, Alderson, Kofler, Sarver, Bolden & Sims, 2008; Martinussen et al., 2005; Willcutt et al., 2005; Hanson, 2011). Two of the meta-analyses compared tasks involving the central executive and tasks not involving the central executive and found that those with ADHD were equally impaired on both types of tasks and that there were no significant differences between their performance on each of these tasks (Martinussen et al., 2005; Hanson, 2011). All of the meta-analyses

compared phonological loop processing and visuospatial sketchpad processing. Martinussen et al. (2005) found that the effect sizes to be in the small to moderate range for tasks involving the phonological loop, but large effect sizes for tasks involving the visuospatial sketchpad. However, the two other meta-analyses did not find a difference between phonological loop and visuospatial sketchpad processing, finding moderate effect sizes for both (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005; Hanson, 2011). Impairments in both the phonological loop and visuospatial sketchpad components of WM are found in children with ADHD, even after controlling for other comorbid conditions, such as dyslexia, language impairments, and other behavioral disorders (Nigg, Blaskey, Huang-Pollock & Rappley, 2002; Kempton, Vance, Maruff, Luk, Costin, & Pantelis, 1999; Hanson, 2011).

Studies have also compared performance between those with ADHD-Combined Type (ADHD-C) and ADHD-Predominately Inattentive Type (ADHD-PI) on WM tasks. A recent meta-analysis found no differences in performance between those with ADHD-C and ADHD-PI on any component on Baddeley's model (Hanson, 2011). This confirms past research, which found similar neuropsychological deficits for these two subtypes of ADHD (Nigg et al., 2002; Nigg, 2006). Additionally, given that inattentive symptoms, but not hyperactive/impulsive symptoms are correlated with WM performance, one would not expect to see differences between the two subtypes of ADHD (Martinussen & Tannock, 2006).

Visuo-spatial WM has also been shown to be able to distinguish between children and adolescents with ADHD from those without ADHD (Westerberg, Hirvikoski,

Forssberg, & Klingberg, 2004). Westerberg et al. (2004) found that visuo-spatial WM distinguished children and adolescents with ADHD from those without better than a continuous performance task, a Go/no go test, and a choice reaction time task. When scores on visuo-spatial WM were taken together with scores on a choice reaction time test, the second best predictor of ADHD status according to Westerberg et al. (2004), they predicted ADHD status with a sensitivity of 74%, which means that it correctly identifies 74% of children with ADHD as having ADHD, and a specificity of 94%, which means that it correctly identifies 94% of children without ADHD as not having ADHD. These predictors have a negative predictive power of 99%, meaning that of the children whom these tests identify as not having ADHD, 99% actually do not have ADHD (Westerberg, et al., 2004). This indicates that there is a low false negative rate. The positive predictive power, which measures the proportion of children identified as having ADHD with these measures, who actually have ADHD, is much lower, at only 19% (Westerberg, et al., 2004). This indicates a high false positive rate. However, this study used a base rate of only 4% to calculate the positive predictive power. In clinical settings, this base rate is likely higher than 4%, which would increase the positive predictive power. For example, if you assume that the base-rate is 20% in clinical settings, then the positive predictive power would be 79%, which indicates a much lower false positive rate. These results seem to suggest that children with ADHD differ in WM abilities from children without ADHD and that WM deficits may be able to be used as a marker of ADHD.

Some hypotheses exist as to why WM deficits are seen in individuals with ADHD. As discussed previously, WM normally develops gradually from preschool through adolescence. However, in one study, this developmentally appropriate increase in WM was not seen to the same degree in children with ADHD (Westerberg, et al., 2004). This provides some evidence that in children with ADHD, the WM system may develop more slowly than children without ADHD. One possible explanation for this delayed development may be brain abnormalities in the areas important for WM functioning. ADHD has been associated with abnormalities in the frontal lobes (Sieg, Gaffney, Preston & Hellings, 1995; Shaw et al., 2007). Shaw et al. (2007) used neuroanatomic magnetic resonance imaging to determine the structure of brains of children and adolescents with and without ADHD. They found that the development in children with ADHD of the cortex, in particular the prefrontal cortex, lagged behind the cortical development of non-ADHD children. Most regions of the prefrontal cortex were more than two years delayed in development in children with ADHD (Shaw et al., 2007). This delayed development of the frontal lobes may be what is contributing to WM deficits in children with ADHD, since WM functions rely heavily on the frontal lobes (Smith & Jonides, 1997).

ADHD has also been found to be a highly heritable disorder, with estimates of 70%-80% heritability (Faraone, Perlis, Doyle, Smoller, Goralnick & Holmgren, 2005). This means that finding genes that may account for ADHD symptomology is particularly important. Often with complex psychological disorders, such as ADHD, it is useful to look for endophenotypes, which are an intermediary between genes and the behaviors

associated with disorders (Jacob & Lesch, 2006). Endophenotypes are thought to be more closely related to the etiological factors of a disease than their diagnostic categories (Morton & Frith, 1995). Endophenotypes are heritable, quantitative traits that indicate a person's likelihood to develop a given disease (Morton & Frith, 1995).

A WM deficit has been proposed as an endophenotype for ADHD (Jacob & Lesch, 2006; Castellanos & Tannock, 2002). One reason that WM has been proposed as an endophenotype is because WM functioning is associated with both prefrontal activity and dopaminergic activity, both of which have also been shown to be correlates of ADHD (Castellanos & Tannock, 2002; Ellis & Nathan, 2001; Smith & Jonides, 1997). WM is also thought to be substantially heritable (Ando, Ono & Wright, 2001; Friedman, Mijake, Young, DeFries, Corley & Hewitt, 2008). As noted earlier, WM performance has been shown to be dependent on dopaminergic modulation of prefrontal neurons (Castellanos & Tannock; Ellis & Nathan, 2001; Braver, Cohen, Nystrom, Jonides, Smith & Noll, 1997). Therefore, variations in genes that affect dopamine functioning, may also affect WM (Castellanos & Tannock, 2002). Many of the genes related to dopamine functioning have also been proposed as candidate genes in ADHD, such as the dopamine transporter gene, DAT1, and the D4 dopamine receptor gene, DRD4 (Li, Sham, Owen & He, 2006; Nigg, 2006). Therefore, frontal lobe activity and DA functioning may help to explain the relationship between WM performance and ADHD.

ADHD is not only associated with poor performance on measures of executive functioning, but also poorer outcomes in many areas of functioning. Children and adolescents with ADHD tend to have more problems with peer acceptance and social

interaction (Bagwell, Molina, Pelham, & Hoza, 2001). Having ADHD may also increase the risk for developing other psychiatric disorders (Smith, Barkley & Shapiro, 2007). Children with ADHD are more likely than children without ADHD to develop oppositional defiant disorder, conduct disorder, and anxiety disorders (Angold, Costello & Erkanli, 1999; Smith, Barkley & Shapiro, 2007). The multimodal treatment study of ADHD found that 68.2% of children with ADHD had at least one comorbid disorder (Jensen, Hinshaw, Kraemer, Lenora, Newcorn & Abikoff et al., 2001). In their sample of 579 children with ADHD, 39.9% also had oppositional defiant disorder, 14.3% had conduct disorder, 38.7% had an anxiety disorder, 10.9% had a tic disorder, and 3.8% had a mood disorder (Jensen et al., 2001). In our previous study, with 46% of participants having oppositional defiant disorder, 39% having an anxiety disorder, and 8% having a mood disorder (Beck et al., 2010).

The symptoms of ADHD normally arise in childhood and frequently ADHD persists into adulthood. One study found that, as reported by parents, ADHD persisted into adulthood in 66% of cases (Barkley, Fischer, Smallish, & Fletcher, 2002). It is estimated that about 4% of the adult, U.S. population are affected by ADHD (Kessler, Adler, Barkley, Biederman, Conners, Delmer, et al., 2006).

Individuals with ADHD often show worse long-term outcomes than their peers. A series of studies by Barkley, Fischer, Smallish, and Fletcher (2006) followed hyperactive children into young adulthood. Although these children were not formally diagnosed with ADHD in childhood, in order to be eligible for the study they needed to fulfill several criteria, which were: to score at least two standard deviations above the

mean on two ratings of hyperactivity, to have parent or teacher reports of poor sustained attention and impulse control and excessive activity level, to have significant behavior problems in several different situations, to have developed their behavior problem before age 6, to have their behavior problem for at least 12 months, and to not have indications of autism, psychosis, thought disorder, epilepsy, gross brain damage, or mental retardation. Given these criteria and their high correlations with ADHD, the authors believe that many of the hyperactive children would have been diagnosed with ADHD according to the DSM-IV. At adolescence, over 70% of the sample met criteria for ADHD according to the Diagnostic and Statistical Manual, third edition, revised. One study found that these young adults were more likely to have lower education performance and attainment, lower employer rated job performance, and based on parent report, fewer close friends, more trouble keeping friends, and more social problems (Barkley, Fischer, Smallish & Fletcher, 2006). Young adults with a history of hyperactivity as children were also more likely to engage in several kinds of socially undesirable behavior, such as early parenthood, being treated for a sexually transmitted disease, being arrested, and drug related activity (Barkley, Fischer, Smallish & Fletcher, 2004; Barkley et al., 2006). These young adults were also more likely to commit crimes such as property theft, disorderly conduct, assault with fists, carrying a concealed weapon, and illegal drug possession (Barkley et al., 2004). Annually, in direct and indirect costs, ADHD is estimated to cost the United States over $30 billions dollars (Birnbaum, Kessler, Lowe, Secnik, Greenberg & Leong et al., 2005). Because of the

persistence of ADHD and its high costs, interventions that lead to permanent or long-term decreases in symptoms of ADHD are needed.

Currently, the most effective stand-alone treatment for the relief of the core ADHD symptoms is medication, specifically psychostimulants (The MTA Cooperative Group, 1999). Children with ADHD who continue to be adherent to their stimulant medications show better treatment outcomes, particularly on core ADHD symptoms, than those who discontinue use and those who continue to take stimulant medications, but are not fully adherent (Charach, Ickowicz & Schachar, 2003; The MTA Cooperative Group, 1999). Reported adherence rates for children and adolescents taking psychostimulant medications range from 35%-100% (Gau, Shen, Chou, Tang, Chiu & Gau, 2006), which provides some evidence that many youth who are prescribed psychostimulants do not fully adhere to them. Psychostimulants are a very effective treatment for ADHD, but their benefits do not persist once they are no longer being taken (Barkley, 2008). Thus, in order to continue receiving the benefits from medications, people with ADHD need to take the medications indefinitely. Non-medication treatments, such as behavioral management and WMT, have been investigated for use with children and adolescents with ADHD.

WM Training (WMT) Research

Research has long supported that practicing an ability leads to improvements in that ability (Ericsson, Krampe & Tesch-Römer, 1993). However, until recently, WM was thought to be a fixed ability, which could not be improved (Engle et al., 1999; Kyllonen

& Christal, 1990; Klingberg, 2010). A previous review of cognitive rehabilitation programs concluded that there was no evidence for training programs being able to improve memory functioning in those with severe memory impairments (Cicerone, Dahlberg, Kalmar, Langebahn, Malec & Bergquist et al., 2000).

However, recent studies have found that WM capacity can be improved by training on WM exercises (Klingberg Fernell, Olesen, Johnson, Gustafsson, Dahlstrom, et al, 2005; Olesen, Westerberg & Klingberg, 2004). One line of research on WMT investigates whether teaching strategies for improving memory could lead to better performance on WM tasks. Specifically, encoding and retrieval strategies have been investigated to see if they lead to improvements in WM. Recent studies have confirmed that the teaching of strategies to aid in memory, leads to better WM performance (Carretti, Borella & De Beni, 2007; Turley-Ames & Whitfield, 2003; Hale, Bookheimer, McGough, Phillips & McCracken, 2007). Researchers have hypothesized that the improvement in strategy use may make the task easier, and, thus, less attention demanding (Engle et al., 1999).

Another line of research investigates whether WM practice alone, without the teaching of strategies, is sufficient to lead to improvements in WM. Garavan, Kelley, Rosen, Rao and Stein (2000) found that WM practice for a total of 2 hours and a total of 8 hours lead to improved response times, but not improved accuracy on WM tasks. However, they reported that ceiling effects on the WM task likely impeded significant improvements in task accuracy. Other studies have found that engaging in increasingly difficult WM tasks over a period of weeks, without the teaching of strategies, is sufficient

in improve WM and performance on a variety of related cognitive tasks (Klingberg et al., 2005; Klingberg, Forssberg, & Westerberg, 2002b).

Klingberg et al. (2002b) first investigated whether WM capacity could be improved in individuals with ADHD by using a computerized WMT program. The WMT exercises included a visuo-spatial WM task, a visuo-spatial version of backwards digit-span, and a spatial-verbal WM task. The WMT program also included an algorithm, which adjusted the difficulty of the training on a trial-by-trial basis to ensure that training was done close to the participant's WM capacity. The training program required participants to perform at least 20 minutes of training 4-6 days a week for at least 5 weeks (Klingberg et al., 2002b). They used a sample of children with ADHD, and half of the children received a control treatment, that did not include the algorithm, so the difficulty was not adjusted and the training was done for less than ten minutes a day. They found that the children who received the full treatment performed significantly better than the control group at post treatment on practiced and unpracticed visuospatial WM tasks.

These results were replicated and expanded on in a randomized controlled trial with children ages 7-12 with ADHD (Klingberg et al., 2005). The WMT in this study included both verbal and visuo-spatial WM exercises similar to those used in the present study. Each training session took about 40 minutes and 20-25 training sessions were done over 5-6 weeks. This study included a placebo treatment like the one explained in the previous study. At post treatment, they found significant improvements of the experimental group on visual spatial WM as measured by Span-board and verbal WM as

measured by Digit-span. Additional improvements were made between post treatment and a three-month follow-up period.

Since Klingberg et al.'s (2002b; 2005) seminal studies several other studies have investigated whether WMT improves WM in individuals with ADHD. Brehmer, Westerberg, and Bäckman (2012) randomly assigned adult participants to the full WMT program described above with an algorithm or to a WMT program without an algorithm that only required participants to remember at most 2 items. They also had participants in the non-adaptive training group complete more trials in order to ensure that they were training for a similar amount of time as the adaptive training group. Brehmer et al. (2012) found that adults in the adaptive training group improved more on WM tasks and self-reported memory problems. Another study compared an adaptive versus non-adaptive WMT program and found that the adaptive training group improved significantly on verbal STM; whereas the non-adaptive training group did not significantly improve on any measure at post-treatment (Van der Molen, Van Luit, Van der Molen, Klugkist & Jongmans, 2010). However, at a 10-week follow-up, the groups had both improved to the same degree on visual and verbal STM and long-term memory and the non-adaptive group improved more on visual-spatial WM. However, this study used a sample of individuals with borderline to impaired intellectual abilities (IQ scores ranging from 55-85) and a relatively low-dose of WMT totaling only 90 minutes of training (6-minute training sessions, three times a week for five weeks), which was not sufficient to improve WM at post-treatment for either group (Van der Molen et al., 2010). Therefore, these study characteristics may account for the findings.

Beck et al. (2010) conducted a study that investigated an intensive, 5-week, home-based WMT similar to that used in the present study. The participants included children and adolescents with ADHD, as verified by parent report as well as a structured clinical interview. Many of the participants also had comorbid diagnoses and were taking psychostimulant medications. When compared to a wait-list control group, there were significant improvements on parent reported working memory. There were no differences in outcome among those with different comorbid diagnoses, ADHD type, or medication status (Beck et al., 2010). These effects were maintained over a 4-month follow-up period. Various other studies find that improvements in WM are stable for 3-18 months after the completion of training (Holmes, Gathercole, Place, Dunning, Hilton, and Elliott, 2010; Klingberg et al., 2005; Dahlin, Nyberg, Bäckman & Neely, 2008; Beck et al., 2010).

School-based WMT for ADHD children has also been investigated and found to improve WM performance (Mezzacappa & Buckner, 2010; Dahlin, 2011). Other studies have looked at WMT that included more gaming elements, such as a story line, animation, identification with a character, and competition (Prins, Dovis, Ponsioen, ten Brink & van der Oord, 2011). When comparing traditional WMT to WMT with gaming elements, children with ADHD who received the training with gaming elements chose to spend more time in training, were more willing to do the training at home, made less mistakes during training, and showed more improvements on an untrained WM task following completion of the training (Prins et al., 2011).

Holmes et al. (2010) assessed the efficacy of stimulant medication and a WMT program for improving symptoms in 25 children between 8 and 11 years of age identified as having ADHD. All children met DSM-IV diagnoses for ADHD, although the procedures for how this was accomplished (e.g., by a structured clinical interview) were not reported. Participants were first measures on phonological loop, visuospatial sketchpad, and central executive components of WM while off medication. They then completed the same measures while taking a psychostimulant medication. After completing these measures, participants continued to take the medication and completed the WMT. Medication alone significantly improved only visuospatial memory performance. When compared to performance after receiving medication alone, WMT lead to additional improvements on all three components of WM. Significant training gains were maintained after 6 months. Many previous studies have documented that psychostimulant medication can improve WM functioning (Holmes, Gathercole, Place, Dunning, Hilton & Elliott, 2010; Barnett, Maruff, Vance, Luk, Costin & Wood et al., 2001; Tannock, Ickowicz & Schachar, 1995; Mehta, Goodyer & Sahakian, 2004; Bedard, Martinussen, Ickowicz & Tannock, 2004). However, WMT leads to additional benefits in WM performance beyond those benefits seen from medication, and WMT appears to lead to benefits among children both on and off psychostimulant medications (Holmes et al., 2010; Beck et al., 2010).

In addition to using WMT with individuals with ADHD, there have been studies of WMT with other clinical and non-clinical samples. In a study of WMT with stroke victims, Westerberg et al. (2007) found that after five weeks of WMT participants

significantly improved on Span board, Digit Span, the Paced Auditory Serial Addition Test, a test requiring WM and control of attention. WMT also leads to improvements in WM among those with multiple sclerosis, acquired brain injuries, adolescents with a history of extremely low birth weight, adolescents with intellectual disabilities, and children with cochlear implants (Lunqvist, Grundström, Samuelsson & Rönnberg, 2010; LØhaugen, Antonsen, Haberg, Gramstad, Vik & Brubakk et al., 2011; Kronenberger, Pisoni, Henning, Colson & Hazzard, 2011; Vogt, Kappos, Calabrese, Stöcklin, Gschwind & Opwis et al., 2009; Van der Molen et al., 2010). In individuals with low WM performance, but not ADHD, an adaptive WMT program, which was heavily taxing on WM, was compared to a control program, which repeatedly presented sequences of two items to remember and, thus, was not taxing on WM (Holmes, Gathercole & Dunning, 2009). Those who received the adaptive training program improved significantly more than those who received the control program on verbal and visuo-spatial WM. These gains were maintained six months after the completion of the training (Holmes et al., 2009).

WMT has been shown to improve WM functioning in typically developing individuals. In four and five year olds visuo-spatial WMT for 15 minutes a day five days a week for five weeks was compared to a control group, which played computer games for the same amount of time (Thorell, Lindqvist, Nutley, Bohlin & Klingberg, 2009). The WMT lead to significantly more improvements on untrained WM tasks. When healthy adults receive intensive, adaptive WMT it leads to improvements on trained and untrained WM tasks (Klingberg et al., 2002; McNab, Varrone, Farde, Jucaite, Bystritsky,

Forssberg et al., 2009; Brehmer, Westerberg, Bellander, Fürth, Karlsson & Bäckman, 2009; Verhaeghen, Cerella & Basak, 2004). Training has also lead to improvement on WM tasks in older adults, with the benefits being maintained for 8-months after treatment (Borella, Carretti, Riboldi & De Beni, 2010).

As discussed previously, earlier training programs to improve memory were generally not successful. Additionally, some more recent attempts to improve WM have not lead to improvements. For example, classroom based strategies, which involved modifying and reducing WM load, encouraging memory-aid strategies, and using direct instruction strategies to improve WM skills, did not lead to any improvements in WM (Elliott, Gathercole, Alloway, Holmes & Kirkwood, 2010). Also, following a 6-week training program involving online, computerized tests of short-term memory, attention, visuospatial processing, and mathematics adult participants did not improve more than a control group on memory tasks (Owen, Hampshire, Grahn, Stenton, Dajani, & Burns et al., 2010). However, subjects in this study widely varied on the number of training sessions completed (1-188) over a six-week period, training sessions were not monitored by study personnel, and each training session only included 6 minutes of training on memory exercises. The difficulty of the training was individualized such that it was challenging for participants. Therefore, this provides some evidence that an individualized training is not sufficient to produce improvements in memory and that memory training may need to occur for a necessary period of time and for a minimum number of sessions for participants to see benefits.

In conclusion, many studies have replicated that WMT can improve WM performance, even on untrained WM tasks for many different individuals, including those with ADHD. This suggests that WMT is not just leading to improvement due to practice, but actually leads to an overall change in WM ability.

WMT Transfer Effects

As discussed previously, WM is important for many other cognitive and academic tasks. Therefore, it stands to reason, that improvements on WM may lead to improvements in other areas of cognitive and academic functioning or to improvements on ADHD symptoms. Since WM has also been implicated to be impaired in individuals with ADHD, studies have begun to investigate whether WMT can lead to improvement in ADHD symptoms.

Studies have found parent reported improvements in ADHD symptoms following WMT (Klingberg et al., 2005; Beck et al., 2010). Improvements in ADHD symptoms are maintained at three to four month follow-up periods (Klingberg et al., 2005; Beck et al., 2010). Johnstone, Roodenrys, Phillips, Watt & Mantz (2010) gave children with ADHD either a training which included an algorithm to keep the training at WM capacity or did not include an algorithm. The training used one type of WMT exercise, which was similar to self-ordered pointing tasks, and one behavioral inhibition exercise, which was a continuous performance task. They found that those who received the training with the algorithm improved more on "significant other" (i.e. non-parent who knows child well such as an aunt/uncle or grandparent) ratings of ADHD symptoms. However, groups

improved equally on measures of parent-reported ADHD symptoms. Another study also did not find parent-reported improvements in ADHD symptoms (Green et al., 2012). When investigating teacher-reported improvements in ADHD symptoms, some studies have not found significant improvements (Klingberg et al., 2005; Beck et al., 2010); however, studies which involved school-based training, have found improvements on teacher ratings of ADHD symptoms (Mezzacappa & Buckner, 2010; Roughan & Hadwig, 2011). Additionally, some studies have found improvements on behavioral measurements of inattention and hyperactivity (Klingberg et al., 2002b; Thorell et al., 2009); however, this is not consistently found (Klingberg et al., 2005). In a recent, placebo-controlled, double-blind study children receiving adaptive compared to non-adaptive training significantly improved on off-task behavior as measured by observing off-task behavior while performing an academic task; however this task was administered in a laboratory setting (Green, Long, Green, Iosif, Dixon & Miller et al., 2012). Individuals with ADHD often show deficits in other executive functions, and WMT has also shown to lead to improvements on executive functions such as inhibition, planning, organization, processing speed, and initiation (Klingberg et al., 2005; Beck et al., 2010; Westerberg et al., 2007; Olesen et al., 2004; Borella et al, 2010).

Some studies have found that after WMT improvements are seen in fluid intelligence (Jaeggi, Buschkuehl, Jonides & Perrig, 2008; Klingberg et al., 2002b; Klingberg et al. 2005; Olesen et al., 2004; Borella et al., 2010; Jaus□ovec & Jaus□ovec, 2012; Schmiedek, Lovden & Lindenberger, 2010; Roughan & Hadwin, 2011). Additionally, Jaeggi et al. (2008) found that as the amount of time spent in WMT

increased, so did the improvements in WM. Also, those who initially had lower fluid intelligence showed more improvement in fluid intelligence following training. One criticism of these studies is that they only use one measure of fluid intelligence, with the exception of Jaus□ovec and Jaus□ovec (2012), Schmiedek et al. (2010), and Roughan & Hadwig (2011). Jaus□ovec and Jaus□ovec (2012) found post-training improvements on Raven's Advanced Progressive Matrices and a spatial rotation task, and borderline statistically significant improvements on a verbal analogies task. Roughan & Hadwin (2011) found improvements on Raven's Standard Progressive Matrices and the Mill Hill Vocabulary Scale; however, these improvements were not maintained at a 3-month follow-up, but this study used a small sample size of only 7 participants receiving training and upon examining the data more closely, the gains in intelligence seen at 3-months were larger than those seen following training. However, the comparison with the control group was not statistically significant. Schmiedek et al. (2010) used Raven's Advanced Progressive Matrices as well as nine tasks from the Berlin Intelligence Structure Test to measure fluid intelligence, which were used to create a latent variable of fluid intelligence. Improvements on this latent variable were seen in younger, but not older, adults following WMT; although, the effect size was small ($d = 0.19$). However, the training program in this study consisted of WMT as well as training of perceptual speed and episodic memory, which may also have contributed to improvements in fluid intelligence. This program also consisted of an average of 100 daily, one hour training sessions, which is more intensive than typical WMT programs. This is consistent with Jaeggi et al. (2008) who found that more intensive training leads to more improvements

in fluid intelligence. Nevertheless, given the use of latent variables, this result is especially encouraging for the idea that WMT leads to an improvement in the overall construct of fluid intelligence.

Other studies have not found improvements in fluid intelligence after WMT (Bergman Nutley, Söderqvist, Bryde, Thorell, Humphreys & Klingberg, 2011; Westerberg et al. 2007; Holmes et al., 2010; Chein & Morrison, 2010; Holmes et al., 2009; Salminen, Strobach & Schubert, 2012). One explanation for these discrepant findings is that WMT improves performance on only certain measures of fluid intelligence or that only certain WMT programs lead to improvements on fluid intelligence. All studies that found improvements in fluid intelligence used tasks similar to Raven's Progressive Matrices to measure fluid intelligence. However, three studies that found improvements in fluid intelligence also used additional measures of fluid intelligence (Jaus□ovec & Jaus□ovec, 2012; Schmiedek et al., 2010; Roughan & Hadwin, 2011). Three of the studies not finding improvements in fluid intelligence used broader measures of intelligence (Bergman et al. 2011; Holmes et al., 2010; Holmes et al., 2009). The others used Raven's Progressive Matrices, but one used participants who had recently had a stroke, which may have impacted the limited improvement seen on this measure (Westerberg et al., 2007) and others used novel WMT programs, which had not been previously studied (Chein & Morrison, 2010; Takeuchi et al., 2011). Interestingly, Salminen et al. (2012) used the same WM task as Jaeggi et al. (2008) and the same outcome measures of Raven's Advanced Progressive Matrices, but one found improvements on this measure and the other did not. Salminen et al. (2012) used 14 days

of training; whereas Jaeggi et al. (2008) used 8, 12, 17, or 19 days of training, with statistically significant improvements on intelligence only seen after 17 or 19 days of training. Therefore, this may demonstrate that a minimum number of sessions are needed for certain transfer effects to occur. Whether transfer effects are seen to fluid intelligence may also be determined by the degree of similarity among the WMT tasks and the fluid reasoning task (Conway & Getz, 2010). In fact, studies have documented that there are strong correlations (as high as r = .90) among WM and fluid intelligence tasks (Kyllonen & Chrystal, 1990; Kane & Engle, 2002). However, this may actually provide further evidence that WMT would be expected to lead to improvements in fluid intelligence since the constructs are highly related.

Few studies have investigated whether WMT can lead to improvements in more applied academic skills. One study did find that school-based WMT lead to improvements on reading comprehension in children with ADHD and learning problems (Dahlin, 2011). Another study found improvements on reading comprehension, as measured by the Nelson– Denny Reading Test, in undergraduate students following WMT (Chein & Morrison, 2010). Loosli, Buschkuehl, Perrig & Jaeggi (2012) found that following WMT typically developing children improved on reading fluency. One study found that following WMT children improved on their accuracy in addition tasks (Witt, 2011). Another study found improved mental arithmetic and comprehension of instructions in children following WMT (St Clair-Thompson, Stevens, Hunt & Bolder, 2010). Holmes et al. (2009) also found that, not immediately after WMT, but 6-months after completion of training children with poor WM had improved on a mathematical

reasoning task. However, other studies have found that following WMT, participants do not improve on word reading or standardized tests of reading or math (Holmes et al., 2009; St Clair-Thompson et al., 2010).

Several researches have suggested that the reason that WMT leads to improvements on other cognitive tasks is that it leads to a general improvement in attention or cognitive control, which is essential for many cognitive tasks (Chein & Morrison, 2010; Jaeggi et al., 2008). Some studies have more directly measured cognitive control, using the Stroop task, and found improvements following WMT (Klingberg et al., 2005; Chein & Morrison, 2010; Takeuchi et al., 2011). If transfer effects are truly the result of changes in ability, then one would expect to find neural mechanisms of change.

Mechanism for Improvement

Investigators have attempted to find what is leading to the improvements in WM abilities and what accounts for transfer effects after training, most of which focus on changes in brain activity. Brain plasticity as a result of practicing a particular ability, has been well documented; however, many previous studies use animal models, focus on structural brain changes, or study changes that occur after a long period (i.e. years) of practice (Rosenzweig & Bennett, 1996). Studies have begun investigating whether brain changes can occur after a relatively short amount of practice on WM abilities. Theorists have proposed that for true brain plasticity to occur there must be both structural and functional brain changes (Lo□vdén, Bäckman, Lindenberger, Shaefer & Schmiedek,

2010). Both structural and functional brain changes have been investigated following WMT.

First, functional brain changes following WMT have been investigated. As discussed previously, developmental increases in brain activity in frontal and parietal regions are associated with increased WM capacity and poor WM performance is associated with decreased activity in these regions (Klingberg et al., 2002; Crone et al., 2006; Klingberg, 2010). Studies have found that after training participants showed increased brain activation in the prefrontal and parietal cortices while performing tasks requiring WM (Olesen, et al., 2004; Westerberg & Klingberg, 2007; Hempel, Giesel, Caraballo, Amann, Meyer & Wüstenberg et al., 2004). Since the changes in activation occurred over several weeks, these authors suggest that skill acquisition, which may involve cortical plasticity, is occurring. Increased activity in the striatum, which has been shown to be activated during WM tasks requiring manipulation of information, has also been found following WMT (Dahlin, Neely, Larsson & Backman, 2008). Dahlin et al. (2008) suggest that in order for WMT to lead to transfer effects that training tasks and the transfer tasks must activate similar brain regions, such that following training, the functional changes in brain activity in those regions is what accounts for the observed transfer effects of improved performance in other cognitive functions. They found evidence of this in their study, which found transfer effects to an untrained task that involved the striatum, but not to another untrained task that does not involve the striatum (Dahlin et al., 2008).

Although some studies have found increased brain activity following WMT, other studies have found decreases in frontal and parietal activity after WMT (Hempel et al., 2004; Schneiders et al., 2011). Petersen et al. (1998) found that after a relatively short amount of practice (i.e. less than 15 minutes) there were reductions in activation of frontal, cingulate, and cerebellum regions. Garavan et al. (2000) found that visual spatial WM tasks activated dorsolateral prefrontal, premotor, cingulate and parietal areas of the brain. This study found that the activation of these regions tended to decrease after both 2 and 8 hours of practice. Gevins, Smith, McEvoy and Yu (1997) found EEG changes in individuals during a WM task after practice, such that less cortical activation appears to be required for the WM tasks after substantial practice. The authors interpret these findings as providing evidence that WMT may increase neural efficiency, thus requiring less brain activation to accomplish the task (Garavan et al., 2000).

Brehmer, Rieckmann, Bellander, Westerberg, Fischer and Backman (2011) gave participants 25 sessions, each taking 25 minutes, of WMT with half of the participants receiving adaptive training and half receiving training at a low WM load. They found that after training participants who received adaptive training showed more decreases in activation in the dorsolateral prefrontal cortex and temporal and occipital regions during a WM task when compared to the group that did not receive adaptive training. Therefore, it appears that adaptive training leads to more increases in neural efficiency, which may account for larger improvement in cognitive tasks after training completion. Another explanation for decreases in functional brain activity is that the practice and improvement of a particular strategy, using particular brain regions, may make other strategies less

used, which then may lead to decreases in brain activation in other areas of the brain which were associated with the previously used strategy that is no longer heavily used (Lo□vdén et al. 2010).

The fact that some studies find increased activation and other find decreased activation in brain activity is not uncommon. In general, neuroimaging studies that involve practice effects have found both increases and decreases in functional activation of the regions believed to be involved in the task (Klingberg, 2010). In a recent review by Klingberg (2010), he suggested that one pattern that emerges is that training for relatively short periods of time (i.e. less than 3 hours) leads to decreases in activation, which may be the result of learning better strategies, and, thus, increased efficiency (Klingberg, 2010). However, when longer periods of WMT are completed and when the WMT also leads to improvements in other cognitive areas of functioning, increases in activation are more common (Klingberg, 2010). Therefore, the increases in activation may be due to improvements in the underlying ability of WM, rather than just practice effects that are limited to WM tasks, which may explain why longer periods of WMT also leads to improvements in other cognitive functions. However, as discussed by Dahlin et al. (2009) there are several studies that do not fit with this pattern and neither whether adaptive versus non-adaptive training is used nor how extensive the training battery is can explain these varying findings in brain activation. In fact, in one study, the exact opposite pattern was found, in that with the same participants, after 2 weeks of training increases in activation in the frontal and parietal lobes, but decreases in activation after 4 weeks of training (Hempel et al., 2004). Clearly, more studies are

needed to clarify when WMT leads to increases, rather than decreases, in brain activity. Nonetheless, these changes in brain activation may represent a biological mechanism to explain why WMT is associated not only with increased WM performance, but also with increased performance on other cognitive tasks and decreased ADHD symptoms.

Due to dopamine's involvement in WM tasks, it has been proposed that there may be changes in dopamine activity after training. McNab et al. (2009) found that after 14 hours of WMT over a five-week period there were changes in dopamine receptor binding in healthy adult males. McNab et al. (2009) used fMRI to identify brain regions that were activated during a WM task. Then, using positron emission tomography (PET), which was conducted during a resting state, the binding potential of D1 and D2 dopamine receptors was determined before and after training in the regions identified to be associated with WM performance. Change in binding potential of the D1 receptors in these regions predicted change ($r^2 = .75$) in WM performance (McNab et al., 2009). Changes in striatal D2 receptor binding potential has also been found using a different WMT program, which emphasized updating more than rote memorization (Ba□ckman, Nyberg, Soveri, Johansson, Andersson & Dahlin et al., 2011). Another study found that those with 9/10-repeat allele of the DAT1 gene showed more improvements in visuospatial WM after training than those with the DAT1 homozygous 10-repeat allele (Brehmer et al., 2009). Another study found that those with a certain variation of LMX1A gene, which is also known to influence DA, showed more improvements on verbal WM following training (Bellander, Brehmer, Westerberg, Karlsson, Fu□rth & Bergman et al., 2011). Clearly there is some evidence that WMT can lead to changes in

the DA system, which may provide a biological explanation for improvements in WM and transfer effects.

Many studies have investigated functional brain changes after WMT, but relatively few have investigated structural changes, which are needed to establish that WMT is leading to true brain plasticity. In terms of structural changes WMT has also been found to lead to increased structural connectivity in parts of the parietal lobes that have white matter and a white matter region adjacent to the corpus callosum (Takeuchi, Sekiguchi, Taki, Yokoyama, Yomogida, Komuro et al., 2010). Furthermore Takeuchi et al. (2010) found that the more training time was associated with more changes in connectivity. Another study by the same group found decreased gray matter volume in regions of the frontal, parietal, and temporal lobes (Takeuchi, Taki, Sassa, Hashizume, Sekiguchi & Fukushima et al., 2011).

Necessary Components of WMT

Studies have begun to investigate the necessary components of WMT. Two factors that have received some attention are whether the training needs to be adaptive and the time spent in training. However, there is not a consensus in the literature on either of these factors. Some studies have found that adaptive training, which adjusts difficulty level based on performance, leads to more improvements than non-adaptive training on some measures (Johnstone et al. 2010; Klingberg et al., 2002b; Klingberg et al., 2005; Holmes et al., 2009); however, others have found no difference between adaptive and non-adaptive training on some measures (Johnstone et al., 2010; Brehmer et al., 2012).

Time in the training program is important to consider when comparing adaptive and non-adaptive trainings. Klingberg et al. (2002) did not ensure that each group engaged in training for the same amount of time and Johnstone et al. (2010) and Holmes et al. (2009) did not specify whether groups trained for the same amount of time, while Brehmer et al. (2012) did ensure both groups were engaged in training for the same amount of time. Brehmer et al. (2012) also used a population of older adults. Both of these differences in study methods may account for these conflicting results. The fact that adaptive training tends to lead to betters results suggests that having training be adaptive may be a necessary component of treatment. As suggested by Jaeggi et al. (2008), adaptive training may be superior to non-adaptive training because it continually engages executive processes, rather than relying on rote memorization or task-specific strategies. As discussed previously, it has been proposed that WMT is leading to brain plasticity, which is proposed as what accounts for the improvement in performance (Klingberg, 2010). In general, for plasticity to occur training must be effortful and engage the participant in tasks at their maximum ability level (Lo□vdén, Bäckman, Lindenberger, Shaefer & Schmiedek, 2010). The use of the algorithm in WMT satisfies these conditions and, therefore, may be necessary for WMT to lead to brain plasticity. This individualization and adaptability of training tasks may be necessary in order for WMT to lead to improvements in other abilities, since this is what is necessary for brain changes to occur (Lo□vdén et al. 2010). Further studies comparing adaptive training and non-adaptive training are needed to definitively determine if having training be adaptive is a necessary component of treatment.

Time spent on WMT is widely variable, ranging from less than 1 hour on one day to over 100 hours over many weeks. To date, no study has reviewed the WMT literature to investigate the relationship between time or days spent on training and the training effects seen; however, one study has attempted to crudely evaluate time and its relationship to effect sizes (Dahlin, Backman, Neely & Nyberg, 2009). This study divided studies into 3 categories based on the magnitude of the effect size found for pre to post-treatment change on the most executively demanding task in the study and then qualitatively compared the training times and concluded that studies with larger effect sizes generally had longer training periods. Jaeggi et al. (2008) gave subjects varying number of training periods (8, 12, 17, or 19 days) and found that those receiving more days of training showed greater improvement. Dahlin et al. (2008) also found a similar pattern by observing increasing effect sizes on trained WM tasks over time, with the effect size after 5 weeks of training being almost 3 times that of the effect size after only 1 week of training.

To date, few studies have compared different training schedules to determine optimum training time or schedules. Vogt et al. (2009) found no difference in outcome when using a more distributed training schedule (two 45-minute training sessions a week for eight weeks) versus a more massed training schedule (four 45-minute training sessions a week for four weeks). However, another study comparing the same training schedules using a sample of healthy adults found that those receiving the more distributed practice improved more than those receiving massed practice on mental speed and some, but not all, measures of STM and WM (Penner, Vogt, Stöcklin, Gschwind, Opwis &

Calabrese, 2012). Jaeggi et al. (2008) administered 25-minute training sessions, five days a week and had participants train for 8, 12, 17, or 19 days. They found that improvements on WM tasks did not differ among the participants that trained for different number of days. However, they found the general trend that as training days increased, so did performance on fluid intelligence. This, and other studies that use relatively short training times, provide support that WM can be improved after a relatively short amount of training; however, for transfer effects to be seen more intensive training may be necessary (Jaeggi et al., 2008).

One study compared visual and auditory WMT programs and found that, the visual training group improved more than the auditory training group on a visual WM task (Schneiders, Opitz, Krick & Mecklinger, 2011). A similar study administered an auditory WMT and found that following training there were only improvements in auditory, but not visual WM (Schneiders, Opitz, Tang, Deng, Xie, Li & Mecklinger, 2012). These two studies suggest that in order to improve overall WM, which is known to consist of verbal and visual components, the WMT must also include verbal and visual components. Training in one modality does not appear sufficient to lead to improvements in the other modality. Clearly, more research is clearly needed to determine the necessary and sufficient components of WM interventions.

Additionally more research is needed to clarify the specific necessary and sufficient components of WMT, especially given the large variation in WMT programs that have been studied.

Review of WM Training

Due to the potentially far-reaching benefits of cognitive training, there have been several recent review articles of WMT and cognitive training programs more broadly (Diamond & Lee, 2010; Morrison & Chein, 2011; Rabipour & Raz, 2012; Shipstead, Redick & Engle, 2012). These reviews have generally found that there is ample evidence that WMT leads to improvements on WM performance. Additionally, compared to other cognitive training programs WMT has more research supporting its efficacy (Diamond & Lee, 2010; Rabipour & Raz, 2012). In terms of transfer effects, the lack of consistently replicated benefits leads reviewers to be more cautious in concluding that WMT can lead to improvements in other areas of functioning; however, reviewers recognize that there is some indication of this and that more research is needed to replicate findings of transfer effects in order to elucidate which areas of functioning are consistently improved by WMT (Diamond & Lee, 2010; Morrison & Chein, 2011; Rabipour & Raz, 2012; Shipstead, Redick & Engle, 2012).

In terms of whether WMT is leading to transfer effects, ultimately, the question is whether WMT is just leading to practice effects or whether it is leading to improvements in the general construct of WM ability or, perhaps even more broadly, whether WMT can lead to improvements in the general construct of cognitive control. If improvements in general abilities were occurring, then transfer effects would be expected. However, if improvements are only seen due to practice effects, but are not leading to an improvement in the underlying ability, then transfer effects would not be expected. Perhaps the most important future direction to answer this question is completion of

studies using latent variables and studies with long-term follow-up because if improvements are leading to overall changes in ability level, then one would expect these benefits to be maintained over time. Additionally, if there is a change in general ability, then there should be underlying neural mechanisms for this change.

Present Study

The present study aims to expand upon previous studies of WMT using an ADHD population. The present study included individuals with ADHD with a variety of comorbid disorder and allowed participants to be taking medication, which is typical of the clinical ADHD population in the United States. Previous studies have found benefit from an adaptive, intensive WMT program when compared to a passive control group and a control WMT, which was not adaptive and involved remembering short spans of items (e.g. Beck et al., 2010; Klingberg et al., 2005). However, as discussed earlier, some previous studies did not control for the amount of time that participants were engaging in the experimental and control WMTs, which the present study does. Previous studies have also found that WMT leads to improvements on parent reports of ADHD symptoms and executive functions (Beck et al., 2010; Klingberg et al., 2005). Some studies have also found improvements on objective measures of WM, inhibition, fluid reasoning, and reading comprehension (Klingberg et al., 2005; Dahlin, 2010). The present study attempts to replicate these findings by investigating whether adaptive WMT leads to more improvements than a non-adaptive WMT on WM. The study also aims to replicate whether there are transfer effects, which lead to improvements in ADHD

symptoms, global executive functioning, and reading comprehension. Additional neuropsychological measures of planning/organization and executive functioning will also be included, which have not been investigated in the past, in an attempt to determine whether WMT can lead to increased performance in these areas. Lastly, the present study will assess the effect sizes of the post-treatment improvements and the clinical significance of these changes. This is particularly important to establish given the high monetary cost and time consuming and effortful nature of the commercially available WMT program used in the present study.

In summary, the present study gave children and adolescents with ADHD WMT, which consisted of 25 sessions lasting 30-45 minutes completed over about 6 weeks. Participants were randomly assigned to either an adaptive WMT program or a control program, which maintained a low-level of difficulty. We predict that the experimental group will show greater improvements on both trained and untrained WM tasks, planning/organization, executive functioning, attention, hyperactivity/impulsivity, and reading comprehension relative to the control group.

Chapter 2: Method

Participants

After obtaining Institutional Review Board approval from Ohio State University, we recruited 60 child and adolescent participants between the ages of 7 and 17 with ADHD. We used several recruitment methods. First, we recruited from a private school intended for children and adolescents with ADHD and/or learning difficulties and located in a large midwestern city. We also recruited participants through other places in the community that provide mental health services: specifically from community mental health centers associated with a children's hospital, the private practice's of two of the principal investigators, and a university psychological services center. Lastly, we put flyers up at the campus of a large midwestern university. All participants had a previous diagnosis of ADHD based on parent-report and a structured clinical interview and questionnaires were administered in order to confirm whether participants meet DSM-IV-TR criteria for ADHD, either ADHD-combined type or ADHD-predominately inattentive type. Twelve participants declined further participation after completion of consent forms and pre-screening measures. Of the 48 participants who were administered the structured clinical interview, 27 met criteria for ADHD-predominately inattentive type and 18 met criteria for ADHD-combined type based on this interview. Based on the clinical interview 7 participants did not meet criteria for ADHD; however, for 6 of them

their parent or teacher rated them as having clinically significant inattentive and/or hyperactive/impulsive symptoms, therefore, they will be included in the analyses. One participant did not have clinically significant ADHD symptoms based on the clinical interview or parent and teacher ratings; therefore, this participant will be excluded from analyses. Children and adolescents were included in the sample regardless of comorbid diagnoses.

Participants were children and adolescents from 7-17 years of age (mean = 12.25 years, 15 females). In terms of ethnicity, 8.3% of participants endorsed Hispanic ethnicity. In terms of race, 81.7% were Caucasian, 6.7% were Hispanic/Latino, 5% were African American, 3.3% were of another race, and 3.3% of participants did not report their race. The modal annual income for families who participated in the study was over $100,000. For additional socioeconomic data, see Table 1. Of the 48 participants who completed the clinical interview, many had disruptive behavior disorders (ODD or CD, 50%), anxiety disorders (39.6%) and mood disorders (6.3%). Of the 48 participants who completed the clinical interview, 45.9% percent presented with no comorbid diagnoses, 25% with one co-morbid diagnosis, 14.6% with two, and 14.7% with three or more co-morbid diagnoses. Of the 48 participants who received pre-treatment measures, 2/3 were taking medication for ADHD. Of the 46 participants who had a completed Social Communication Questionnaire, Lifetime Form (SCQ), 13% fell above the cutoff of 15, which is indicative of a possible Autism Spectrum Disorder.

Of the 60 participants to agree to participate in the study, 12 did not complete any pre-data collection due to the decision not to participate in the study either due to lack of

time or motivation, leaving 48 participants who were allocated to the experimental or control group. Three participants completed the pre-treatment data collection but never began the training program, leaving 45 participants who began the training program. Of these 45 participants, 10 did not complete the training program due to the following reasons: 2 due to frustration with the training program, 5 due to being unmotivated to complete the program, 2 due to family dysfunction, and 1 due to illness. Thus, 27.08% of participants who were allocated to a treatment, did not complete the treatment. Of the 35 participants who completed the training program, 6 could not be scheduled for a post-treatment data collection meeting, leaving 29 participants whom have complete pre-treatment and post-treatment data. This means that of those who were randomly assigned to the experimental or control group, 39.58% of them dropped out of the study before completing it. Of those who completed pre-treatment data (and thus were randomly assigned to a group), 7 of the 25 who were randomly assigned to the experimental group dropped out and 12 of the 23 who were randomly assigned to the control group dropped out.

One of these participants who completed post-treatment data did not clearly have a diagnosis of ADHD based on clinical interview and rating scales, and, thus, was excluded form the analyses. This left 17 participants in the experimental group and 11 participants in the control group who completed post-treatment measures.

Measures

Parent/Teacher Report Measures

Demographic Questionnaire. The attached demographic questionnaire (Appendix A) was used to assess age, grade, gender, socioeconomic status, race, and ethnicity.

DSM-IV ADHD Symptoms Checklist. This measure asked parents whether they endorse each of the 18 ADHD symptoms from the DSM-IV-TR for their child (Appendix B).

The Social Communication Questionnaire, Lifetime Form (SCQ). A parent of the participants completed the SCQ to screen for Autism Spectrum Disorders (ASDs). The SCQ uses a cut-off score, with scores of 15 or greater being indicative of a possible ASD. The standardization sample of the SCQ included a clinical sample of 160 individuals with ASDs and 40 individuals with other psychiatric diagnoses (Rutter, Bailey & Lord, 2003). The SCQ was found to effectively differentiate between those with ASDs and those without. With a cutoff of 15, the SCQ has a sensitivity of .85, specificity of .75, positive predictive value of .93, and negative predictive value of .55 when distinguishing between those with versus those without an ASD (Rutter et al., 2003). The subtests of the SCQ highly correlated (.73-.89) with the corresponding subtests of the Autism Diagnostic Interview-Revised (ADI-R), which is considered one of the gold standards in diagnostic assessment of ASDs (Rutter et al., 2003). The total scores of the SCQ and ADI-R also highly correlated (r = .78) (Rutter et al., 2003). The internal consistency is acceptable for all age ranges (alpha = .84-.93).

Children's Interview for Psychiatric Syndromes-Parent Version (P-ChIPS). One of two graduate students administered the P-ChIPS in its entirety to determine each participant's diagnoses on 20 Axis I disorders (See Table 2). The P-ChIPS is based on the DSM-IV and screens for a variety of disorders (Weller, Weller, Rooney, & Fristad, 1999). The child version of this measure, which asks the same questions as the parent version except that they are rephrased to reflect the reporting source (i.e. "do you…" is changed to "does your son/daughter…"), has been shown to be a reliable and valid diagnostic instrument in clinical research for children and adolescents from 6-18 years of age (Fristad, Glickman, Verducci, Teare, Weller, & Weller, 1998; Fristad, Cummins, Verducci, Teare, Weller, & Weller, 1998; Teare, Fristad, Weller, Weller, & Salmon, 1998a, 1998b). Diagnoses of children obtained by administering the P-ChIPS to parents correlate moderately with diagnoses obtained using the child version of this interview (Fristad, Teare, Weller, Weller, & Salmon, 1998). The P-ChIPS also correlates moderately with clinicians' diagnoses of children, with a 76% agreement for ADHD (Fristad, Teare, et al., 1998). Fristad, Teare, et al. (1998) also found the P-ChIPS to have average sensitivity of 87% and an average specificity of 76%, with a 100% sensitivity and 44% specificity for ADHD.

Conners 3 Parent and Teacher. We administered both the parent and teacher versions of the Conners 3, Short Form, which is the revision of the Conners' Rating Scale-Revised, Short Form. The Conners measures are widely used in clinical and research settings for assessing ADHD symptoms (Connors, Sitarenios, Parker, & Epstein, 1998a; 1998b). The Conners 3 has six subscales: Inattention, Hyperactivity/Impulsivity,

Learning Problems, Executive Functioning, Defiance/Aggression, and Peer Relations. These subscales were revised based on exploratory and confirmatory factor analyses. The standardization sample for the Conners 3 was similar to the 2000 census in the make up of race and ethnicity and included 50 males and 50 females in each age range. There is good internal consistency, with Cranach's alpha above 0.90 for all parent and teacher subscales (Arffa, 2010). The test-retest reliability ranged from .82-.98 for the parent form and .78-.90 for the teacher form. Discriminant validity was established, with the Conners 3 successfully discriminating between clinical and non-clinical samples as well as among different clinical groups including: ADHD, disruptive behavioral disorders, and learning disorders (Arffa, 2010). There was also evidence for convergent validity, in that there were significant correlations between the Conners 3 and the Behavioral Assessment System for Children, 2nd Edition, the Achenbach System of Empirically Based Assessment, and the Behavioral Rating Inventory of Executive Functioning. The mean correlation between the parent and the teacher version of the Conners 3 was .60 (Arffa, 2010).

Behavior Rating Inventory of Executive Function (BRIEF) Parent and Teacher Forms. We administered the parent and teacher forms of the Behavior Rating Inventory of Executive Function (BRIEF, Gioia, Isquith, Guy, & Kenworth, 2000a). It is intended to measure executive functioning in children and adolescents from 5-18 years of age and consists of eight subscales, which are Emotional Control, Inhibit, Initiate, Monitor, Organization of Materials, Plan/Organize, Shift, and Working Memory. In addition, the BRIEF includes two index scores, which combine various subscales into one

scale. The Behavioral Regulation Index is comprised of the Inhibit, Shift, and Emotional Control subscales. The Metacognition Index is comprised of the Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor scales. Finally, the BRIEF includes the Global Executive Composite, which is a summary score that includes all eight subscales. Reliability studies show high internal consistency, with Cronbach's alpha between 0.80 and 0.98 and test-retest reliability ranging from 0.79-0.88 over a two-week period (Gioia, Isquith, Guy, & Kenworth, 2000b). Convergent validity has been established with other measures of inattention, impulsivity, and learning skills in clinical ADHD populations (Gioia, et al., 2000b).

Participant-Administered Measures

Automated Working Memory Assessment (AWMA). The short form of the AWMA was administered which includes four subtests, digit recall, listening recall, dot matrix, and spatial recall, which measure verbal STM, verbal WM, visuospatial STM, and visuospatial WM, respectively. In children aged 4-11 over a four week period the test retest reliabilities for the four subtests used in the present study, digit recall, listening recall, dot matrix, and spatial span were, .84, .81, .83, and .82 respectively (Alloway et al., 2006). In individuals with low WM abilities the test-retest reliability is low to moderate (r = .27-.50) when a test interval of about 9 months was used (Alloway, Gathercole, Kirkwood & Elliott, 2008). When using the AWMA and the WISC-IV to assign children to either an average or a low WM group, the results from the AWMA agreed with results fro the Working Memory Index of the WISC-IV at a rate of 89.3%.

Agreement was 89.3% for the AWMA and WISC-IV digit span subtest and 71.4% for the WISC-IV letter-number sequencing subtest. Because of the relatively few studies investigating the psychometric properties of the AWMA, the WISC-IV, which has well known and adequate psychometric properties, will also be used to assess WM.

Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). The digit span and letter-number sequencing (LNS) subtests of the WISC-IV were administered in order to assess working memory. These scores were also used to calculate the Working Memory Index (WMI). The WISC-IV standardization sample, which was representative of the U.S. population in terms of age, gender, race, ethnicity, and parent education level according to the 2000 census data, consisted of 2,200 children and teenagers (Wechsler, 2003). The internal consistency reliabilities of digit span, LNS, and the WMI ranged from .81-.92, .85-.92, and .90-.93, respectively. The test-retest reliability over a period of on average 32 days is .81 for digit span, .75 for LNS, and .85 for the WMI (Wechsler, 2003). Exploratory and confirmatory factor analyses were conducted, which verified the structure of the WISC-IV. Validity of these WM subtest was demonstrated by a correlation of .74 between the WMI and the Children's Memory Scale Attention/Concentration subtest. The WMI also correlates with the Freedom From Distractibility Factor from the WISC-III (r = .74; Wechsler, 2003).

Children's Color Trails Test (CCTT). The CCTT was administered in order to assess overall attention and executive functioning. The CCTT is thought to involve cognitive flexibility, executive skills, psychomotor speed, sequencing, and visual attention (Llorente, Williams, Satz & D'Elia, 2003). The CCTT has two parts (1 & 2).

In both parts the child is presented with a page with circles of two different colors with numbers in them. In CCTT-1 the child is presented with a page with 15 circles with the numbers 1-15 in them and is asked to connect the numbers in order as quickly as possible. All even numbers are presented in one color and all odd numbers in the other color. In CCTT-2, there are 30 circles numbered 1-15, with each number being presented twice, once in each of the colors. The child is again instructed to connect the circles in order as quickly as possible, but is also asked to alternate between each of the colors, so that he/she has to not only choose the appropriate number, but also the appropriate color when completing the task.

The CCTT standardization sample consisted of 678 healthy children between the ages of 8-16 years. The test-retest reliability in a sample of children aged 6-12 years with ADHD without any comorbid diagnoses was calculated for overall completion time on the CCTT-1 and the CCTT-2. The correlation was moderate for the CCTT-1 ($r = .46$) and for the CCTT-2 ($r = .66$) with two months between testing and for the CCTT-1 ($r = .68$) and CCTT-2 ($r = .60$) with four months between testing (Llorente, Voigt, Berretta, Rennie, Fraley, & Satz, P. et al., 2002). When calculating the reliability of the clinical interpretation of the CCTT-1 and the CCTT-2 based on completion time, the reliability was high ($r = .90$-$.99$) for two and four month intervals (Llorente et al., 2002). Therefore, if a child was determined to be impaired at the first administration, there was a high level of agreement at the second administration. Increasing age was associated with decreasing completion time in the standardization sample (Llorente et al., 2003). This provides some evidence of the validity of the CCTT being sensitive to neurological

functioning because as children age their neurological functioning is also improving. Performance on the CCTT-1 and CCTT-2 are also moderately to highly correlated with the similar, Children's Trail Making Test A & B (r = .74 and .67, respectively). The CCTT is also moderately correlated (r = .35-.51) with the Test of Variables of Attention (Llorente et al., 2002).

The CCTT is able to distinguish between those with neurological problems and those without and appears to be sensitive to frontal lobe dysfunction (Williams, Rickert, Hogan, Zolten, Satz, D'Elia, et al., 1995). Studies have found differences between ADHD individuals and normal controls on the CCTT (Williams et al., 1995).

Tower of London Drexel University: 2nd Edition (TOL). The TOL was used to assess planning and organization. Since planning and organization are thought to be somewhat dependent on WM, WM is also likely involved in the TOL (Levin, Fletcher, Kufera, Harward, Lilly & Mendelsohn et al., 1996; Welsh, Satterlee-Cartmell & Stine, 1999). The task involves moving beads along pegs in order to make a particular patter in as few moves as possible. The TOL includes a Total Moves Score, which were the primary outcome for this study. Other scores included on the TOL are total correct, rule violations, time violations, initiation time, execution time, and total time.

The standardization sample included 1,234 children and adults, including 244 children with ADHD. A study of test-retest reliability with a sample of ADHD children aged 7-10 years used a test-retest interval of 5-92 days. The reliability for total moves was high (r = .80; Culbertson & Zillmer, 1999). In an ADHD sample the TOL is significantly, although weakly to moderately, correlated with other measures of executive

functioning including the Wisconsin Card Sort Test, Selective Reminding Test, Stroop Color-Word Test, and Rey Osterrieth Complex Figure (Culbertson & Zillmer, 1999). The tower of London is moderately correlated (r = .55) with the Porteus Mazes, another measure of planning (Krikorian, Bartok & Gay, 1994).

Children with ADHD have been found to score lower on the TOL total moves score than normal controls (Cornoldi, Barbieri, Gaiani, & Zocchi, 1999). When using the TOL to distinguish between children with and without ADHD the sensitivity is .76, specificity is .81, positive predictive value is .73, and negative predictive value is .77.

Woodcock-Johnson III Normative Update Tests of Achievement (WJ-III-NU). The following subtests of the WJ-III-NU were administered: story recall (immediate and delayed), understanding directions, passage comprehension, and reading vocabulary. The standardization sample included 4,740 school-aged individuals from 100 U.S. communities, which were representative of the U.S. population in terms of race, ethnicity, and parent education level (Woodcock, McGrew, Schrank, & Mather, 2007). The split-half reliabilities ranged from 0.79-0.89 for story recall, 0.62-0.85 for understanding directions, .73-.96 for passage comprehension, .74-.88 for story recall-delayed, and .82-.93 for reading vocabulary. The test re-test reliability for a test interval of less than a year was only available for passage comprehension and was .91. For a test re-test interval of one year, the reliabilities ranged from .73-.89 for passage comprehension and.53-.62 for story recall (Woodcock et al., 2007).

Gray Oral Reading Tests, Fourth Edition (GORT-4). The GORT-4 was used to assess reading comprehension. It requires participants to read a passage aloud to the

examiner and then answer questions about the passage, without being able to refer back to the passage. The GORT-4 has two forms: A & B. Form A was given at pre-treatment and form B was given at post-treatment. The GORT-4 was normed on a sample of 1,677 individuals (Wiederholt & Bryant, 2001). The internal consistency of the GORT is good, with a Cronbach's alpha of .98 for comprehension with a sample of individuals with ADHD. The GORT-4 also has a small standard error of measurement (SEM = 1) for comprehension. The correlation between Form A and Form B ranges from .71-.86 for comprehension. When Form A and Form B were administered two weeks apart, the correlation was adequate (.78-.95; Wiederholt & Bryant, 2001).

Summary

To summarize, the specific outcome variables that were examined in the present study include:

1. Working Memory as measured by:
    - Working Memory Index (WMI) from the WISC-IV
    - Automated Working Memory Assessment (AWMA) Digit Recall
    - AWMA Dot Matrix
    - AWMA Listening Recall
    - AWMA Spatial Recall
    - BRIEF Parent Working Memory
    - BRIEF Teacher Working Memory
2. Attention as measured by:

- Conners Parent Inattention
- Conners Teacher Inattention
- Woodcock-Johnson Oral Language

3. Hyperactivity/Impulsivity as measured by:
    - Conners Parent Hyperactivity/Impulsivity
    - Conners Teacher Hyperactivity/Impulsivity
4. Reading Comprehension as measured by:
    - Woodcock-Johnson Reading Comprehension
    - Grey Oral Reading Test Reading Comprehension
5. Planning/Organization as measured by:
    - Tower of London Move Score
    - Tower of London Execution Time
    - BRIEF Parent Planning/Organization T-score
    - BRIEF Teacher Planning/Organization T-score
6. Other executive functions as measured by:
    - Children's Color Trails Test Interference Index
    - Conners Parent Executive Functioning
    - BRIEF Parent Global Executive Control (GEC)
    - BRIEF Teacher Global Executive Control (GEC)

Because all of the objective measures of working memory were highly correlated (all p-values < .01), a working memory composite score was created by averaging scores

on the WMI, AWMA Digit Recall, AWMA Dot Matrix, AWMA Listening Recall, and AWMA Spatial Recall.

Procedures

For those participants recruited from the private school or private practices school officials or clinicians gave parents of children and adolescents with ADHD or attention difficulties a flyer telling them about the study. If interested, study personnel gave parents information about the requirements to participate in the study. Parents who chose to participate in the study then filled out consent forms for themselves and their children. They also filled out the Demographic Questionnaire and DSM-IV ADHD Symptoms Checklist.

For participants who were recruited elsewhere, they were instructed on the flyer to call or e-mail study personnel. A phone screening was then conducted in which the study requirements were explained to parents and the DSM-IV ADHD Symptoms Checklist was administered over the phone to determine presence of ADHD symptoms.

Parents who appeared to have a child with ADHD and who wished to participate were then contacted to schedule a pre-treatment data collection meeting on the OSU campus. At the pre-treatment meeting parents were administered the P-ChIPS, Conners-3, BRIEF, and SCQ. Parents also completed a semi-structured interview, which asked about past diagnoses, past psychiatric test results, learning problems, medication status, vision or hearing problems, and any history of seizures or tics. Parents were asked to try not to change medication or other courses of treatment over the training period. Parents

were then trained on the implementation of the Cogmed WMT program, given instructions for how to download and log into the program, given an opportunity to ask questions about the program, and a time was scheduled for the researchers to contact them once they had started the training program. During this time, the participants completed the subtests of the WISC-IV and WJ-III, the CCTT, the TOL, the AWMA, and the GORT-4, in that order. Trained undergraduate and graduate students administered these measures. Participants were given breaks as needed and rewarded for their efforts with small prizes (i.e. candy, stickers, super balls, etc). Before beginning the training program, one of the participant's teachers completed the Conners-3 and BRIEF.

We then assigned participants to either the experimental or control group by using a random numbers table. The experimental group received the Cogmed WMT program that is available commercially and those in the control group received a modified WMT program that was less taxing on WM, only requiring participants to remember, at most, 5 items. In both groups the WMT consisted of a computer based training program that participants did in their home under the supervision of one parent. The training included 25 sessions completed in about 6 weeks, with each session taking 30-45 minutes. Each session included fifteen trials of eight of a possible twelve WM exercises (Table 3).

If a participant took less than 30 minutes to complete a session, the number of trials in each exercise was increased until the training time was at least 30 minutes. In the experimental group, the training included an algorithm that continually increased or decreased the difficulty of each exercise according to the child's performance, so the

participants were always working at or near their WM capacity. This training is the commercially available Cogmed RM, which is similar to that used in past studies (Beck et al., 2010; Westerberg et al., 2007; Olesen et al., 2004; Klingberg et al., 2005; Klingberg, Forssberg & Westerberg, 2002). The control group received a program with a modified algorithm, such that their maximum training level was five items and when they got an item wrong, they were dropped down to only having to remember two items. The difficulty also increased at a much slower rate. Therefore, most of their time was spent training at a low level of WM load (less than 5 items).

In both groups, a trained experimenter viewed the results of each session, spoke with the participant and their parent about the quality of their sessions that week, and discussed any problems. The other purpose of these calls was to provide positive reinforcement to the participants for continuing with the training program. The participants also received rewards from their parents for doing the training. These rewards varied by participant and the participant and their parent decided on them before the training began. They included things such as getting $1 for each training day completed or getting to pick something fun to do for every week of training completed.

One month after completion of the 25 sessions, participants and their parents attended the post-treatment meeting, where the same assessments were completed as the pre-treatment meeting, except that the parents did not complete the P-ChIPS, SCQ, or semi-structured interview. The parents were asked about any changes in medication or any other treatments during the training period as well as any big changes in the child's

life during the training period. The participants' teachers also completed the Conners-3 and BRIEF at this time.

Data Analysis

Preliminary t-tests and a non-parametrics test, Fisher's Exact Test, (which serves the same purpose as a chi-squared test) were conducted to determine whether there were any pre-treatment differences between the experimental and control groups as well any differences between those who completed post-treatment measures and those who did not on age, sex, ethnicity, which experimenter was their coach, comorbid diagnoses, subtype of ADHD, ADHD medication status, or any pre-treatment outcome measures. Because there were more than two groups, chi-squared analyses were used to compare groups on income and race.

Several types of statistical analyses were used to determine whether there were significant pre-treatment to post-treatment changes on any outcome measures and to determine if the experimental group improved significantly more than the control group on the outcome measures. First, repeated-measures analysis of variance (rANOVA) was used. As discussed in Gueorguieva & Krystal (2004), rANOVA continues to be the most common method used for analyzing repeated-measures data. The effect size, partial eta-squared, was used to determine the magnitude of the effect. A partial eta-squared of .01 will be considered small, .06 with be considered moderate, and .14 will be considered large (Cohen, 1988).

Although rANOVA is the most widely used method, analysis of covariance (ANCOVA) of post-treatment measures with pre-treatment measures as a covariate is a more preferred method (Van Breukelen, 2006). This is due to the fact that ANCOVA has more power than rANOVA. However, ANCOVA is biased when there are group differences at baseline, such as those that may be seen in studies that do not use random assignment (Van Breukelen, 2006). Because the present study uses random assignment, which should eliminate pre-treatment group differences, ANCOVA analyses will also be conducted. As discussed by Raab, Day & Sales (2000), which covariates are included when using ANCOVA should be decided before beginning data analysis and should not be data driven. Because in our previous study (Beck et al., 2011) there were no significant correlations between any demographic variables and outcome variables, no additional covariates were included.

There are assumptions that are made when using ANOVA, namely that data is normally distributed and that there is homogeneity of variance across time and groups. These assumptions were tested and some variables violated these assumptions. Therefore, non-parametric analyses, which do not require these assumptions, were also completed. Specifically, the Wilcoxon Signed Rank Test was used. Because non-parametric tests for investigating interactions in repeated-measures data are not well established, non-parametric tests were only be used to investigate the pre-treatment to post-treatment change, without evaluating the differences between the experimental and control group.

Lastly, there are benefits to using hierarchical linear modeling (HLM) to analyze repeated measure data. HLM has more flexibility than ANOVA, which often makes it the preferred method for repeated measures analysis (Gueorguieva & Krystal, 2004). HLM analyses were conducted to evaluate whether there were differences between groups in the degree of change on outcome measures from pre-treatment to post-treatment. One advantage of HLM is that it does not require excluding participants who have missing outcome data and, thus, has greater power than ANOVA analyses and does not have as biased of results due to the exclusion of those who are missing data. Also, there was some variation in the time between when pre-treatment and post-treatment measures were completed. HLM allows one to take into account non-uniform time periods between measurements. In order to account for these variations, time was calculated individually for each participant in that the date that participants completed the pre-treatment measure was considered time point zero and the other time point was determined by the number of weeks after time point zero when post-treatment data was collected.

Statistically significant results were evaluated for clinical significance. We evaluated clinically significant change by determining if the post-treatment scores of each participant are closer to the mean of the normative or non-normative population on each measure, following the method described by Jacobson & Truax (1991). A cutoff point was established by averaging the mean of the normative population and the mean of the ADHD population for each measure. Post-treatment scores that fall below this cutoff point were classified as showing clinically significant change.

Next, we calculated a reliable change index (RCI), also described by Jacobson & Truax (1991), which determines whether changes in participants' scores are statistically reliable, or unlikely due to measurement error. The RCI is defined as:

$$RCI = \frac{\text{pre treatment score} - \text{post treatment score}}{\text{standard error of the difference between the two scores}}$$

The standard error of the difference between the two scores can be defined as:

$$\sqrt{2}\ (\text{standard error of measurement})^2$$

The standard error of measurement can be defined as:

$$\text{Standard deviation at baseline} \sqrt{(1 - \text{test-retest reliability of the measure})}$$

An RCI of greater than 1.96 is significant at the $p = .05$ level of significance. Therefore, subjects with an RCI of greater than 1.96 were classified as showing reliable change.

Lastly, we used a clinical equivalence test described by Kendall, Marrs-Garcia, Nath, and Sheldrick (1999). This test attempts to determine if after treatment the participants are clinically equivalent to the normative population. If at post treatment the participants were equivalent to the normal population, then one would expect the differences between the post treatment mean and the mean of individuals in the non-clinical population to be zero. A two-sided t-test was conducted comparing the means at post-treatment and the normative population mean (i.e. a Standard Score of 100, T-score

of 50, or Scaled Score of 10), with the null hypothesis that these two means were equal. A non-significant result of this test, would lead to the failure to reject the null hypothesis that the means were equal, thus providing support that at post-treatment, the participants had normalized. Clinical equivalence tests were only conducted on measures where, at pre-treatment, participants scored significantly worse than the normal population. This was determined by calculating the t-tests described above with pre-treatment data, with a statistically significant result indicating that the participants were significantly different than the normal population at pre-treatment. Due to the small sample size, clinical equivalence tests were not conducted with teacher measures.

Chapter 3: Results

Preliminary Analyses

Multiple Comparisons

Because we are making multiple comparisons, our alpha level will be adjusted. The Sidak-Bonferroni correction, with α_{FW} being the familywise error rate and c being the number of comparisons, is (Keppel & Wickens, 2004):

$$\alpha = 1 - (1 - \alpha_{FW})^{1/c}$$

In addition, because the adjusted alpha levels are quite small, all analyses that have a moderate or large effect size are also noted.

Power Analysis

A power analysis indicated that to be able to detect a moderate effect (partial eta squared = .06) with a power of .80, we need a sample size of 34 participants when using rANOVA for the interaction between time and group. Our actual power to detect the group by time interaction with a sample of 28 was 0.72. This means that there is not enough power to detect a moderate effect size. Therefore, effect sizes that are moderate or large will be reported, regardless of significance level.

Differences at baseline between the experimental and control groups

Based on Fisher's Exact Test and chi-squared analyses, the experimental and waitlist control groups do not differ at baseline on their sex, income, ethnicity, race, medication status, coach, ADHD type (Inattentive or Combined), presence of comorbid disorders, presence of internalizing disorders (Mood and Anxiety Disorders), presence of externalizing disorders (Oppositional Defiant Disorder or Conduct Disorder), or possibility of an ASD, with all *p*-values > .05 (Tables 3 and 4). The experimental and control groups were also compared on all pre-treatment outcome measures as well as Conners' parent and teacher defiance/aggression and learning problems, scores on the Social Communication Questionnaire (SCQ), total number of comorbid diagnoses, age, and number of ADHD inattentive and hyperactive/impulsive symptoms. Using a family-wise alpha of .05 and 23 comparisons, the Sidak-Bonferroni corrected p-value used to determine significance is .0022, and there were no significant differences using this alpha level. However, several comparisons had a p-value of .05 or less including the Automated Working Memory Assessment (AWMA) Spatial Recall subtest, t (40) = 2.10, p = .04, the Conners Parent Learning Problems subtest, t (45) = -2.42, p = .02, and the Woodcock-Johnson Reading Comprehension subtest, t (44) = 2.05, p = .05.

When using only the participants who completed post-treatment measures, there were no significant differences using the Sidak-Bonferroni corrected p-value. However, several comparisons had a p-value of .05 or less including BRIEF Parent Planning/Organization, t(26) = -2.32, p = .03, BRIEF Parent WM, t(26) = -2.25, p = .03, Conners Parent Executive Functioning, t(26) = -2.33, p = .03, and Conners Parent

Inattention, $t(26) = -2.20$, $p = .04$, such that the control group scored worse than the experimental group on these measures. The control group also had a higher number of externalizing disorders, $t(26) = -2.79$, $p = .01$.

Differences at baseline between completers and non-completers of post-treatment

Those who dropped out of the study and those who did not were compared on pre-treatment variables and it was found that those who dropped out and those who completed the study did not differ on gender, ethnicity, race, medication status, coach, ADHD type (Inattentive or Combined), presence of comorbid disorders, presence of internalizing disorders (Mood and Anxiety Disorders), presence of externalizing disorders (Oppositional Defiant Disorder or Conduct Disorder), or possibility of an ASD, with all *p*-values > .05. Those who completed the study and those who dropped out were also compared on all pre-treatment measures, total number of comorbid diagnoses, age, and number of ADHD inattentive and hyperactive/impulsive symptoms using the Sidak-Bonferroni corrected p-value of .002. Using this p-value there were no significant difference between those who did versus did not complete the study. However, some comparisons had a p-value of .05 or less including the Conners Parent Defiance Aggression subscale, $t(45) = -2.34$, $p = .02$ and the Tower of London Move Score, $t(43) = 2.12$, $p = .04$, with the non-completers scoring worse on these measures.

Correlations of the measures

Pre-treatment scores on outcome variables, Conners' parent and teacher defiance/aggression and learning problems, scores on the Social Communication Questionnaire (SCQ), total number of comorbid diagnoses, age, and number of ADHD inattentive and hyperactive/impulsive symptoms were each correlated with pre to post-treatment change scores on all outcome variables. Using the Sidak-Bonferroni corrected *p*-value of .005, there were several significant correlations, which are reported in Table 9 along with all correlations that have *p*-values of .05 or lower. As would be predicted, for many variables higher scores at pre-treatment were associated with more improvements on the measure at post-treatment, which may represent regression to the mean. Higher levels of parent-rated defiance/aggression were associated with more improvement on parent-rated WM, inattention, and global executive functioning. Higher level of learning problems at pre-treatment was associated with smaller improvements on spatial STM, but more improvement on teacher reported planning/organization and global executive functioning. A higher number of comorbid diagnoses was associated with fewer improvements on parent-rated inattention.

Clinical Equivalence Tests

Several one sample t-tests were conducted to determine whether the participant scored significantly worse than the normal population on measures at pre-treatment, with the null hypothesis that participants scores were equal to the mean score in the normal population (i.e. a T-score of 50, Standard Score of 100, or Scaled Score of 10).

Participants scores were significantly worse than the normal population at pretreatment on the AWMA Dot Matrix t (25) = -2.43, p = .02, Working Memory Index, t (26) = -3.58, p < .01, BRIEF Parent WM, t (27) = 9.89, p < .01, Conners Parent Inattention, t (27) = 13.81, p < .01, Conners Parent Hyperactivity/Impulsivity, t (27) = 5.37, p < .01, BRIEF Parent Planning/Organization, t (27) = 8.26, p < .01, BRIEF Parent Global Executive Control, t (27) = 8.00, p < .01, and Conners Parent Executive Functioning, t (27) = 7.23, p < .01. On all other parent-report and participant-administered measures, participants' scores were no different from the normal population.

Post-Treatment Performance on Outcome Measures

Both the change from pre-treatment to post-treatment and the group (experimental or control) by time (pre-treatment or post-treatment) interaction were examined for statistical significance. The ANOVA assumptions of normality and homogeneity of variance were tested. These assumptions were violated for several measures. Measures that did not have a normal distribution included the BRIEF Parent WM, BRIEF Teacher Planning/Organization, Conners Parent Hyperactivity/Impulsivity, Conners Teacher Hyperactivity/Impulsivity, TOL Execution Time, and TOL Move Score. Measures that did not have equality of variances include the AWMA Spatial Recall, W-J Oral Language, and TOL Move Score. Therefore, non-parametric techniques were used in addition to ANOVA to analyze the data.

Working Memory

A Sidak-Bonferroni corrected *p*-value of .006 was used to evaluate statistical significance. Using this *p*-value, no measures had a significant group-by-time interaction based on rANOVA, ANCOVA, and HLM analyses (Tables 10 and 11). However, as discussed previously, our study was underpowered. The WISC-IV WMI had a large effect size and 82.4% of the experimental group met the cut-off for clinically significant change, with 41.2% of participants showing reliable change on the WMI. The AWMA dot matrix had a moderate to large effect size with 70.6% of participants in the experimental group meeting criteria for clinically significant change and 47.1% showing reliable change (Table 22). The BRIEF Teacher WM subscale also had a large effect size, but it was in the opposite direction in that the experimental group got significantly worse than the control group following treatment (Tables 10 and 11). At post-treatment, participants in the experimental group, t (16) = 1.29, p = .22, and control group, t (10) = -0.84, p = .42, did not score significantly different from the normal population (i.e. a Standard Score of 100) on the AWMA Dot Matrix. On the WMI, participants in the experimental group, t (16) = 1.36, p = .19, and control group, t (9) = -2.18, p = .06, did not score significantly different than the normal population. However, participants in the experimental group, t (16) = 5.03, p < .01, and control group, t (10) = 6.04, p < .01, still scored significantly worse than the normal population on BRIEF Parent WM at post-treatment.

When evaluating whether there was a significant improvement from pre-treatment to post-treatment, regardless of group membership, participants significantly improved on

AWMA spatial recall, WISC-IV WMI, the WM Composite Score, and BRIEF Parent WM based on rANOVA, and these improvements had large effect sizes (Tables 10 and 12). The same results were found using the Wilcoxon Signed Rank Test and HLM except that using these measures there was also a significant improvement on AWMA dot matrix with a large effect size. Additionally the change in AWMA digit recall had a moderate effect size. See Table 22 for the percentage of participants who reached the cutoff used for clinically significant and reliable change. There was also large effect size for the change in BRIEF Teacher WM, although in the opposite direction as predicted (Tables 10 and 12). Clinical equivalence was evaluated for the WMI, AWMA Dot Matrix, and BRIEF Parent WM since at pre-treatment participants' scores were significantly worse the normal population on these measures.

ADHD Symptoms

A Sidak-Bonferroni corrected p-value of .017 for measures of inattention and .025 for measures of hyperactivity/impulsivity were used to evaluate statistical significance. No measures that had a significant group-by-time interaction. However, as discussed previously, our study was underpowered. There was a large effect size for Conners Parent Inattention, but the control group showed more improvement than the experimental group. There was also a large effect size for Conners Teacher Inattention and Conners Teacher Hyperactivity/Impulsivity, but with the experimental group being substantially worse than the control group at post-treatment. There was a moderate effect size for Conners Parent Hyperactivity/Impulsivity, with the control group showing

more improvement than the experimental group (Tables 13 and 14). At post-treatment on Conners Parent Inattention participants in the experimental group, t (16) = 7.11, p < .01, and the control group, t (10) = 6.96, p < .01, scored significantly worse than the normal population (i.e. a T-score significantly higher than 50). Participants in the experimental group, t (16) = 7.97, p < .01, and the control group, t (10) = 2.65, p = .02, scored significantly worse than the normal population on Conners Hyperactivity/Impulsivity at post-treatment.

When evaluating whether there was a significant improvement from pre-treatment to post-treatment, regardless of group membership, participants significantly improved on Conners Parent Inattention with a large effect size (Tables 13 and 15). The same result was found using the Wilcoxon Signed Rank Test and HLM. However, only about 17% of participants showed clinically significant change (i.e. a post-treatment T-score of below 60) and only about 28% of changes were reliable (i.e. unlikely due to error) (Table 22). Additionally the change in Conners Teacher Inattention had a large effect size and Conners Teacher Hyperactivity/Impulsivity had a moderate effect size, but with these effects in the opposite direction (i.e. being worse at post-treatment). There was also a moderate effect size for improvement on W-J Oral Language (Tables 13 and 15).

Reading Comprehension

A Sidak-Bonferroni corrected p -value of .025 was used to evaluate statistical significance. No measures that had a significant group-by-time interaction based on rANOVA, ANCOVA, or HLM (Tables 16 and 17). When evaluating whether there was

a significant improvement from pre-treatment to post-treatment, regardless of group membership, there was a statically significant improvement on W-J Reading Comprehension based on rANOVA and HLM, but not the Wilcoxon Signed Rank Test (Tables 16 and 18). On W-J Reading Comprehension about 59% of participants showed clinically significant change, and this change was reliable in about 27% of participants (Table 22). There was a large effect size for improvement on W-J Reading Comprehension and a moderate effect size for improvement on GORT Comprehension (Tables 16 and 18).

Executive Functions

A Sidak-Bonferroni corrected *p*-value of .013 was used for measures of planning/organization and executive functioning to evaluate statistical significance. No measures that had a significant group-by-time interaction. There was a large effect size for TOL Execution Time, but only based on ANCOVA analyses and not based on rANOVA or HLM (Tables 19 and 20). About 81% of participants in the experimental group showed clinically significant change with about 47% showing reliable change (Table 22). A moderate effect was found for the CCTT Interference Index, but only based on rANOVA, and not on ANCOVA. There was also a large effect size for Brief Teacher Planning/Organization based on rANOVA and ANCOVA, but with the control group doing notably better than the experimental group at post-treatment (Table 19 and 20). At post-treatment on BRIEF Parent Planning/Organization participants in the experimental group, $t\,(16) = 3.66$, $p < .01$, and the control group, $t\,(10) = 5.64$, $p < .01$,

scored significantly worse than the normal population (i.e. a T-score significantly higher than 50). Participants in the experimental group, t (16) = 4.76, p < .01, and the control group, t (10) = 6.04, p = <. 01, scored significantly worse than the normal population on BRIEF Parent Global Executive Control at post-treatment. On Conners Parent Executive Functioning participants in the experimental group, t (16) = 3.39, p < .01, and the control group, t (10) = 6.03, p < .01, scored significantly worse than the normal population.

When evaluating whether there was a significant improvement from pre-treatment to post-treatment, regardless of group membership, participants significantly improved on TOL Execution Time, BRIEF Parent Planning/Organization, and BRIEF Parent GEC, with large effect sizes (Table 19 and 21). The same results were found using the Wilcoxon Signed Rank Test, rANOVA, and HLM. On the TOL Execution time, about 74% of participants showed clinically significant change, and this change was reliable for 36% of participants. On BRIEF Parent Planning/Organization about 32% of participants showed clinically significant change, but only about 14% of participants showed reliable change. On the BRIEF Parent GEC about 36% of participants showed clinically significant change, but only about 14% had reliable change (Table 22). Additionally the change in BRIEF Teacher Planning/Organization and BRIEF Teacher GEC had a large effect size, but with participants being worse at post-treatment. The Conners Parent Executive Functioning had a moderate effect size for improvement from pre-treatment to post-treatment (Table 19 and 21).

Chapter 4: Discussion

The present study aimed to expand upon the current WMT literature by determining whether a difficult, adaptive WMT led to more improvements on ADHD symptoms and executive functions than an easier WMT, which required participants to remember fewer items but took the same amount of time to complete. The study used a clinical population of children and adolescents with ADHD, many of whom were taking ADHD medication and had comorbid diagnoses. The study used both participant administered academic and neuropsychological measures as well as parent and teacher report measures. In general, participants improved on most objective and parent-report measures of executive functioning and WM; however, those who received the adaptive training did not tend to improve more than those who received the easier training. This lends some support for the idea that WMT can lead to improvements in broad cognitive functions; however, future studies need to investigate what the necessary components of WMT are and whether the improvements following WMT are clinically significant and not just due to practice effects, rater expectancy effects, or regression to the mean.

Preliminary Analyses

Despite using random assignment of the participants who completed post-treatment data, using a significance level of .05, there were significant differences at pre-

treatment on parent rated executive functions (specifically planning/organization and WM) and parent rated inattention, with the control group scoring worse on these measures. The control group also had a significantly higher number of externalizing disorders. One may think that these differences were due to the fact that many participants dropped out. In fact, when random assignment was initially used and the entire sample (not just those who completed post-treatment data) was used for comparisons between the experimental and control groups, there were no differences on any of these measures. However, there were differences on other variables (spatial recall, learning problems, and reading comprehension). It is unclear why, even with random assignment, there were still some pre-treatment differences. One explanation may be the sheer number of comparisons conducted, in that if you compare two groups on enough variables, finding at least some differences is likely. Thus, when a corrected p-value was used to account for these many comparisons, there were no statistically significant differences between groups. Nonetheless, when interpreting the results, these pre-treatment differences between the experimental and control group should be considered.

Due to the high rate of dropout our study was underpowered. Adjusting the alpha level to account for multiple comparisons, also has the effect of further decreasing power. Therefore, we had a high probability of making a type II error (i.e. concluding that there is not a significant effect when in fact there is one). Thus, when interpreting results there will be less reliance on statistical significance and more reliance on effect sizes, which are less reliant on sample size.

Our high level of dropout is inconsistent with previous studies using WMT with individuals with ADHD, which report dropout rates ranging from about 2-17% (Beck et al., 2010; Klingberg et al., 2005; Mezzacappa & Buckner, 2010; Dahlin, 2011; Prins et al., 2011; Johnstone et al., 2010; Green et al., 2012; Thorell et al., 2009). In order to help elucidate what factors made participants more likely to drop out of the study, those who completed the study were compared to those who did not on all demographic and pre-treatment variables. Result indicated that those who did not complete the study had higher levels of parent-rated defiance/aggression and performed worse on the TOL task, indicating possible poorer planning/organization skills. It is unclear what made the drop out rate of our study higher than previous studies using WMT. The possible causes of increased drop out rate in our study could be possible at each step in the study. First, the reason why some participants declined participation after completing screening measures and signing consent forms was that they viewed the time commitment as too great. Secondly, in addition to the time spent on WMT (which is generally uniform across all Cogmed WMT studies) our study required coming to the university for two two-and-a-half to three-hour meetings for pre and post-treatment data collection. We also were unable to provide monetary incentives for participants to complete the training program and to complete post-treatment data collection, which may have contributed to dropout at these time points.

There also may have been something about the control group training program, which has not been used previously, which may have contributed to dropout rates. Perhaps the control WMT was too boring or frustrating for participants to complete.

Although our control treatment was similar to control treatments used in previous studies (e.g. Klingberg et al., 2005), it was somewhat different in that participants were required to engage in the control treatment for a longer period of time, which may have made the training more boring. Also, the control training did have an algorithm to adjust difficulty based on performance. The difficulty of items increased much slower when participants got items correct, and they were only permitted to reach a maximum of having to remember five items. When participants got one item wrong, the difficulty level dropped down to the lowest level. Participants in our study may have seen this control training as more frustrating than previously used control trainings, and perhaps combined with the parent-rated defiance of participants, the children and adolescents in the study may have simply refused to continue. In retrospect, a pilot study using the control treatment should have been conducted in order to determine whether it was a viable condition for participants to complete. A previous study instructed participants in a control training to engage in speed training in order to enhance motivation (Brehmer et al., 2012). Perhaps if this had been done with our participants it may have enhanced motivation and decreased dropout rate.

Discussion of Outcome Measures

Teacher Report Measures

Due to a bulk of teacher measures not being received in the mail (despite school personnel reporting that they were sent in one package), few teacher measures were

available for analysis. Because the sample is very small, it is difficult to draw any generalizations from this data. Therefore, no conclusions will be drawn from this data.

WM Outcomes

There were pre-treatment to post-treatment improvements on all WM measures except for the AWMA Listening Recall, which is a measure of verbal WM that is very different from the training tasks. Not finding improvements on some WM measures, is inconsistent with some previous studies of WM training, which find improvements on all measures of WM following training, including untrained tasks (Diamond & Lee, 2010; Rabipour & Raz, 2012; Shipstead et al., 2012; Morrison & Chein, 2011). However, Holmes et al. (2009) also did not find improvements on some WM subtests of the AWMA, including dot matrix and another visuospatial WM subtest that was not used in the present study. However, this study did not use listening recall. Holmes et al. (2009) also found that there were not improvements on a verbal WM task. Holmes et al. (2010) also used the AWMA and did not find post-treatment improvements on the Digit Recall and Dot Matrix subtests as well as two other subtests that tap verbal WM. However, when participants were re-assessed at follow-up, improvements were seen on all measures of WM. One other study did not find improvements on a verbal STM tasks (Bergman Nutley et al., 2011).

The findings of Holmes et al. (2010) and other studies that have found additional gains at follow-up (Beck et al., 2010; Klingberg et al., 2005) suggest that perhaps it may take longer for gains to be seen. Another possibility is that the Cogmed WMT only

improves certain areas of WM and that previous studies, which have used outcome measures that are very similar to the training, have failed to realize the more limited nature of the improvements in WM.

The effect size for the interaction of group and time, such that the experimental group improved more than the control group, was large for the measure of spatial short-term memory and for a measure of verbal working memory. However, on all the other measures of WM, including the WM composite, there were not differences between the experimental and control group at post-treatment. Thus, the experimental training, which is commercially available, had minimal benefits over the control training, which was not as taxing on WM. This is inconsistent with previous studies that have found little benefit of the non-adaptive training. It is important to note that that control training in the present study was adaptive, but the algorithm for adjusting difficulty limited the number of items that participants were required to remember and, on average, participants were required to remember fewer items than in the WMT used as the experimental treatment. The purpose of the algorithm used in the experimental training is to keep participants doing most of their training at their WM limit, which, clearly, the control treatment did not achieve (Klingberg et al., 2005). In conclusion, although we used an active control WMT that included some elements of WMT, it did not keep participants training at their WM limit, which has been proposed as a key element of WMT (Jaeggi et al., 2008).

Additionally, as discussed earlier, many of the previous studies comparing adaptive and non-adaptive trainings did not ensure that participants in each group were training for the same amount of time (Klingberg et al., 2005; Holmes et al., 2009;

Johnstone et al., 2010). It is important to note the one study that did ensure equal training times, did not find differences between the adaptive and non-adaptive training (Brehmer et al., 2012). Therefore, whether the training is adaptive or not may not be as important as the amount of time that is spent in the training. To date, there is no definitive study which determines the amount of WMT that is needed in order to lead to benefits, but studies generally suggest that the more training that is completed, the more benefits that are seen from training (e.g., Dahlin et al., 2009; Dahlin et al., 2008; Jaeggi et al., 2008). Therefore, it is possible that the reason why previous studies found greater benefit from the adaptive training than the non-adaptive training was because participants spend more time completing the adaptive training.

It is also possible that, similar to Holmes et al. (2010), more time following training is needed in order to see differences between experimental and control trainings. Perhaps if we had conducted a follow-up assessment a few months after completion of training, we may have seen more group differences. However, there is one study that is inconsistent with this, in that both the adaptive and non-adaptive training groups show more gains at follow-up (Van der Molen et al., 2010).

Nonetheless, in the present study on one measure of non-verbal STM and one measure of verbal WM participants did improve more when they received the adaptive training as compared to the non-adaptive training. Therefore, this lends some evidence for the hypothesis that training WM with an adaptive training that keeps participants training at their WM limit is necessary. There have also been theoretical arguments for the justification of the necessity of the algorithm, such as the algorithm being needed in

order to see transfer effects because it forces participants to engage in more executive processes rather than rote memorization, which would not be expected to lead to improvements on other tasks (Jaeggi et al., 2008). Along these lines, others have proposed that WMT is leading to brain plasticity, which is what accounts for the improvements in performance, and that for plasticity to occur participants need to be engaging effortfully in the training task (Lovden et al., 2010; Klingberg, 2010).

In terms of the effect sizes seen, the moderate to large effect sizes found in the present study are generally consistent with those found in other WMT studies (Diamond & Lee, 2010). Effect sizes of this magnitude are generally considered to lend evidence that the treatment is effective. However, there is limited data regarding whether the WMT leads to clinically significant changes in functioning following treatment or real world benefits that can be observed by those interacting with participants. Jacobson & Truax (1991) proposed a way to quantify a clinically significant change, by considering the change clinically significant if, at post-treatment, a participant is closer to the mean of the normal population than to the mean of the disordered population. This conceptualization can be useful for measures where there is a large difference among those in the two populations (i.e. the Conners Inattention and Hyperactivity/Impulsivity subscales). However, on measures of WM, those with ADHD generally do somewhat worse, but not substantially worse, than the general population. This means that many participants were already closer to the normal population mean at pre-treatment and that relatively small changes (i.e. 3 standard score points) could be enough to allow a participant to reach the cutoff (See Table 22). Therefore, for the measures of WM, we

generally do see that the majority of participants are classified as having clinically significant change, but when examining where they started at pre-treatment, the change in the percent of participants meeting the cutoff is minimal for most measures. Thus, for measures of WM perhaps this way of conceptualizing clinically significant change is not as useful. On a measure of non-verbal STM (Dot Matrix) and verbal WM (WMI), participants initially scored significantly worse than the normal population (i.e. a Standard Score less than 100). However, at post-treatment participants in both the experimental and control groups did not score statistically significantly worse than a standard score of 100. This provides some evidence for clinically significant change in that participants normalized on these measures.

It is especially important to consider whether the change was reliable or likely due to error (i.e. improvements just because of re-testing or extraneous factors in the testing environment that lead to improvements). Ideally, collecting follow-up data can be useful for determining whether the change is reliable. However, since that was not possible for the present study, a reliable change index was used instead. For the Dot Matrix, WMI, and parent report of WM about one-third of participants showed reliable change, which is similar to a previous study and consistent with the idea that, at least for some participants, change is maintained over longer periods of time, as evidenced by sustained improvements at follow-up points months after the completing of the WMT (Beck et al., 2010). However, for the Spatial Recall test, few participants showed reliable change, which may indicate that this improvement is not stable and largely due to error.

ADHD Symptom Outcomes

There were significant improvements in parent-reported symptoms of inattention from pre-treatment to post-treatment. However, contrary to expectation, there were more improvements in the control group than in the experimental group. This may be due to the fact that the control group started higher at pre-treatment than the experimental group, so more regression to the mean would be expected in the control group. When the groups were combined, there was a large effect size for the improvement in parent-reported inattentive symptoms, which is consistent with previous studies (Beck et al., 2010; Klingberg et al., 2005). However, in the Klingberg et al. (2005) study only participants who received adaptive training improved on parent-rated inattention. Johnstone et al. (2010) also found that when family members, other than parents, rated participants on ADHD symptoms after treatment, participants who received adaptive training were rated as more improved than those who did not. These two findings are inconsistent with our study in that both groups improved. However, Johnstone et al. (2010) did find that both participants receiving adaptive and non-adaptive training improved equally on parent-rated inattention. Therefore, similar to the discussion above, it is difficult to determine whether adaptive training is a necessary component of treatment in order to lead to improvements in parent-reported inattention. Also, when combining groups and examining pre to post-treatment improvements, there is no longer a control group for comparison, so it is possible that the improvement seen is no more than what would be seen in a group of participants who had not received any treatment completing the measures again. However, our previous study did use a waitlist control group and there

was a significant difference between the control group and the active treatment group, which makes this less likely (Beck et al., 2010). The reliable change index also takes into account test-retest reliability, so can be used as a gage of for how many participants the change is unlikely due to completing the measure again. The changes in inattention were reliable in a little over one-quarter of participants, which is consistent with our previous study, but is also quire low (Beck et al. 2010). This indicates that for the majority of participants the change seen was more likely due to error than to stable changes due to the treatment program.

There has yet to be a study where a blind rater reports improvements in individuals with ADHD following training. In all of the previous studies, in which significant changes on rating scales were seen, the raters were aware that the participants were receiving the treatment program (Beck et al., 2010; Klingberg et al., 2005; Johnstone et al, 2010; Mezzacappa & Buckner, 2010; Roughan & Hadwig, 2011). Therefore, the reported improvements may be due to the placebo effect. However, in the Klingberg et al. (2005) study raters were not aware what group participants were in and rated the experimental group as significantly more improved than the control group on ADHD symptoms, so the placebo effect did not come in to play in that study. Also, Beck et al. (2010) found a trend for teacher rated improvements in WM and teachers were not aware of whether their students were receiving the WMT or not. There also have been some studies using more objective measures of attention and hyperactivity, which have found improvements following training (Klingberg et al., 2002b; Green et al., 2012). This lends more evidence to the fact that WM training actually is leading to

improvements above and beyond the improvement that is seen based on expectation alone. In terms of how much of an improvement was seen, the cutoff for what was considered clinically significant change was if the participants post-treatment score was below a t-score of 60. Only about 18% of participants achieved this endpoint, compared to about 4% who had scores below 60 at pre-treatment. Participants also did not normalize on measures of parent-reported ADHD symptoms, in that at post-treatment they were still rated as significantly higher than average (i.e. a T-score of 50). Therefore, although WMT does appear to reduce parent-reported symptoms of inattention, it is unclear how meaningful these changes are clinically and whether the changes are large enough to justify the effort needed to complete the training program.

Reading Comprehension Outcomes

Both the experimental and control group improved on reading comprehension tasks following treatment, but neither group improved more than the other. This is consistent with other studies, which have found improvements in reading comprehension following training (Dahlin, 2011; Chein & Morrison, 2010). However, because groups were combined for pre to post-treatment analyses, there was no control group, so it is possible that these improvements are due to the expectation of improvement on the part of the participants and their parents or to practice effects from having completed the measure previously. Both of the previous studies investigating improvements in reading comprehension had an inactive control group, which eliminates the possibility of improvements only being due to practice effects, but does not eliminate the possibility of

the effects being due to the placebo effect (Dahlin, 2011; Chein & Morrison, 2010). Therefore, additional, placebo controlled studies are needed to determine whether improvements in reading comprehension following WMT are due to treatment factors specific to WMT. A little over one-fourth of participants showed reliable change and no follow-up studies to date have investigated whether these improvements in reading comprehension are stable. Therefore, more research is needed to determine whether these improvements are stable over time.

Executive Functioning Outcomes

There were no significant group by time interactions in that the experimental group improved more than the control group on any executive functioning outcomes. Significant pre to post-treatment improvements were seen on most measures of executive functioning included planning/organization as measured by the TOL execution time and by parent report and overall executive functioning as rated by parents. The one exception was the Children's Color Trail Test (CCTT). For the trail-making test, there was a trend that the control group improved, while the experimental group got worse. However, when the groups were combined, there was no significant improvement from pre to post-treatment. A previous study found no improvements on a trail-making task following WMT (Takeuchi et al., 2011). This may be because trail-making tasks involve a wide variety of cognitive functions, including fine motor skills, motor speed, visual scanning, and processing speed, which are not as closely related to WM.

Findings of improvements on parent-rated planning/organization and broad executive functioning measures are consistent with our previous study (Beck et al., 2010). About one third of participants surpassed the cut-off for clinically significant change at post-treatment. However, at pre-treatment about 20-25% of participants had surpassed this cutoff. Additionally, at post-treatment participants still scored significantly worse than the normal population on these measures. Therefore, there is not strong evidence for clinically significant improvements. This is further supported by the fact that only about 14% of participants showed reliable change on these measures.

No previous study had investigated whether improvements in planning/organization could be seen on neuropsychological tests. The improvements on the TOL suggest that WMT may be promising for improving performance on measures of planning/organization. On the TOL execution time almost three-quarters of participants surpassed the cutoff for clinically significant change at post-treatment, compared to less than half who had reached this cutoff at pre-treatment. Almost half of the experimental group participants showed reliable change, compared to one-fifth of the control group. However, participants did not show significant improvement on the TOL move score, which is sometimes regarded as the primary outcome measure for the TOL. Thus, further exploration is needed regarding whether WMT leads to measureable improvements in planning/organization and whether these improvements carry over into real world tasks (i.e. improved organization at home or school).

The fact that WMT led to improvements on many different academic and cognitive measures offers support for the idea that WM is important for many cognitive

tasks and that improvement in WM could then lead to improvement in broad cognitive functions.

Limitations

One of the major limitations of the study was the high rate of dropout. This is inconsistent with previous studies using this WMT program (e.g. Beck et al., 2010). It is likely that the higher demand put on participants in this study compared to previous studies, namely being required to attend two 2.5-3 hour meetings at the university, contributed to dropout. In the future it would appear beneficial to try to decrease dropout rates. In order to do so, it may be necessary to decrease the length of the meeting times. This could be accomplished by allowing some measures to be completed over the phone (i.e. parent interview) or be completed and then mailed in (i.e. questionnaire). It also may prove beneficial to decrease the number of measures given to participants during the meeting, especially given that some of the measures measured the same construct. Although there were reasons for including each measure, in retrospect, it may have been better to include just one measure of each construct, even if that meant including some measures with less than ideal psychometric properties. Additionally, participants and there parents could have been offered an incentive for completing the pre-treatment and post-treatment data collection meetings, which may have decreased dropout rates.

In addition, because more participants in the control group dropped out than in the experimental group, it is also possible that the control treatment used was more demanding than anticipated or that it lead to more boredom or frustration than the

experimental treatment. If the same control treatment is used in the future, it should be piloted with a group of participants and participants and their parents should be interviewed to determine possible reasons why it appears to be associated with higher dropout rates. The control treatment could then be modified accordingly.

It is possible that participants' tendency towards defiance/aggression and poor planning and organization contributed to dropout rates since those who dropped out had higher parent-reported defiance/aggression and lower planning/organization abilities, as measured by the Tower of London. Being more selective of participants to include in the study could reduce dropout. Because dropout was associated with higher defiance/aggression and lower planning/organization, participants with higher levels of defiance/aggression and lower levels of planning/organization may not be as good of candidates for WMT. It is also possible that there are other characteristics of participants whom are likely to dropout, which were not measured in the current study. Based on experimenter observations, families whom dropped out of the study appeared to have more familial dysfunction and stress, so it may be beneficial to include a measure of this in future studies, which may be used as a prescreening measure to determine eligibility for the study. However, increasing exclusion criteria will make it more difficult to recruit eligible participants.

Lastly, when combing the groups to examine pre-treatment to post-treatment improvements there is no inactive control group for comparison. Therefore, it is possible that the improvements were due to practice effects or to the placebo effect.

Future Directions

Future studies are needed to further investigate the necessary and sufficient components of WMT and to investigate whether certain features of training are more likely to lead to benefits in certain areas. Specifically, it should be further investigated whether adaptive training is needed and whether a WMT that keeps participants training at their WM limit is needed for certain gains. From the results of the present study and after reviewing the WMT literature, it is unclear whether either of these components of WMT are truly a necessary component of the treatment. The amount of time spent in WMT and the training schedule should also be further studied to determine what the minimum amount of time needed is to see certain training gains and to determine if certain training schedules (i.e. one day a week for 2 hours versus 5 days a week for 30 minutes) lead to more benefits. Studies have also used many different types of WM exercises in training, with some studies using only one WM exercises and other studies using a wide variety of exercises. Therefore, future studies should investigate which WMT exercises are necessary and sufficient to lead to improvements.

Further assessment of whether improvements seen as a result of training are clinically significant is needed. Unfortunately, clinically significance is often difficult to evaluate. It is often difficult to determine how much change needs to be seen on a standardized measure in order to indicate that there has been some observable, functional change in performance. The inclusion of more applied, academic measures, such as grades or report card scores, appears promising for determining whether there are real world changes observed following treatment. Another option would be to gather more

qualitative data from participants, their families, and their teachers about what improvements, if any, they have observed. However, these measures generally lack good psychometric properties. There does not appear to be a clear answer as to how to best assess clinical significance and it will likely take a combination of methods to determine whether this treatment is leading to observable improvements in the lives of the individuals who partake in the treatment. Along these lines, additional investigation of whether the benefits outweigh the costs of this treatment is also important. The treatment is time consuming for families and comes at a monetary price when used in clinical practice. Therefore, a future goal should be to determine whether the benefits seen are large enough to justify the cost, in terms of effort, time, and money. Finding a metric for measuring this may prove difficult, but it is an important question to better understand, especially when discussing the use of the program in a clinical setting.

There also needs to be further assessment of what specific cognitive functions are improved following WMT. There is generally not a consensus about which areas consistently improve following WMT. Some studies may find improvements in one function, but then other studies do not find improvement in this function. Therefore, the findings of WMT studies need to be replicated in order to determine which improvements are consistently found and to determine whether certain features or length of WMT are associated with greater improvement in certain functions.

Conclusion

There were generally few interactions in that the experimental group improved more than the control group, with the exceptions of a measure of spatial short-term memory and a measure of verbal working memory. This may indicate that there is little added benefit to having a WMT that keeps participants working at their WM capacity over a training that requires remembering fewer items. It may be that the amount of time spent in the training program is a more salient factor. However, our study has several limitations, most notably high dropout rates, which lead to low power. Therefore, with a larger sample size and less bias due to a truly random design, perhaps the adaptive training group would have outperformed the non-adaptive training group.

Generally, there were pre-treatment to post-treatment improvements on many measures including WM, inattention, reading comprehension, planning/organization, and executive functioning. This provides some evidence that WMT may be promising for improving a broad range of cognitive and academic functioning. This is consistent with the theory that WM is important for a broad range of cognitive functions. Therefore, if WMT were leading to true improvements in the construct of WM, then one would expect to see improvements in other functions that rely on WM. However, the lack of comparison to an inactive control group and limited amount of clinically significant and reliable change makes it unclear how substantial the improvements in these areas are.

References

Alloway, T. P., Gathercole, S. E., Kirkwood, H. & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology, 28,* 725-734.

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.

Ando, J., Ono, Y. & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavioral Genetics, 31*, 615–624.

Arffa, S. (2010). Review of the Conners 3rd. In R.A. Spies, K.F. Geisinger, and J.F. Carlson (Eds.), *The eighteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from Mental Measurements Yearbook with Tests in Print database.

Angold, A., Costello, E. J., Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40,* 57-87.

Bäckman, L., Karlsson, S., Fischer, H., Karlsson, P., Brehmer, Y., Rieckmann, A., et al. (2011). Dopamine D1 receptors and age differences in brain activation during working memory. *Neurobiology of Aging, 32*, 1849-1856.

Ba□ckman, L., Nyberg, L., Soveri, A., Johansson, J., Andersson, M. & Dahlin, E. et al. (2011). Effects of working-memory training on striatal dopamine release. *Science, 333*, 718.

Badcock, J. C., Michie, P. T. & Rock, D. (2005). Spatial working memory and planning ability: Contrasts between schizophrenia and bipolar I disorder. *Cortex, 41,* 753-763.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417–422.

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

Baddeley, A. D, & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Baddeley, A. D. & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8,* 485-493.

Bagwell, C. L., Molina, B. S. G., Pelham, W. E., & Hoza, B. (2001). Attention-deficit hyperactivity disorder and problems in peer relations: Predictions from childhood to adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1285–1292.

Barkley, R.A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121:1,* 65-94.

Barkley, R.A. (2006). *Attention-deficit/hyperactivity disorder*. In E.J. Mash & R.A. Barkley (Eds.), Behavioral and Emotional Disorders in Adolescents. (pp. 63-112). New York: Guilford Press.

Barkley, R.A. (2008). Global issues related to attention-deficit/hyperactivity disorder from childhood to young adulthood. *Postgraduate Medicine, 120:3,* 48-59.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111,* 279–289.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2004). Young adult follow-up of hyperactive children: Antisocial activities and drug use. *Journal of Child Psychology and Psychiatry*, 45:2, 195–211.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2006). Young adult outcomes of hyperactive children: Adaptive functioning in major life activities. *Journal of the American Academy of Child and Adolescent Psychiatry, 45:2,* 192-202.

Barkley, R.A., Murphy, K. & Kwasnik, D. (1996). Psychological adjustment and adaptive impairments in young adults with ADHD. *Journal of Attention Disorders, 1,* 41-54.

Barnett, R., Maruff, P., Vance, A., Luk, E. S. L., Costin, J., Wood et al. (2001). Abnormal execution function in attention deficit hyperactivity disorder: The effect of stimulant medication and age on spatial working memory. *Psychological Medicine, 31,* 1107-1115.

Bedard, A. C., Martinussen, R., Ickowicz, A. & Tannock, R. (2004). Methylphenidate improves visual-spatial memory in children with attention deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 260–268.

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

Bellander, M., Brehmer, Y., Westerberg, H., Karlsson, S., Fu□rth, D. & Bergman, O. et al. (2011). Preliminary evidence that allelic variation in the LMX1A gene influences training-related working memory improvement. *Neuropsychologia, 49*, 1938-1942.

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.

Birnbaum, H.G., Kessler, R.C., Lowe, S.W., Secnik, K., Greenberg, P.E., Leong, S.A., et al. (2005). Costs of attention-deficit-hyperactivity disorder (ADHD) in the US: Excess costs of persons with ADHD and their family members in 2000. *Current Medical Research and Opinion, 21 (2)*, 195-205.

Borella, E., Carretti, B., Riboldi, F. & De Beni, R. (2010). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging, 25*, 767-778.

Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., & Noll, D.C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage, 5,* 49-62.

Brehmer, Y., Rieckmann, 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.

Brehmer, Y., Westerberg, H. & Bäckman, L. (2012). Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Frontiers in Human Neuroscience, 27,* 1-7.

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.

Cain, K., Oakhill, J., & Bryant, P. (2000). Phonological skills and comprehension failure: A test of the phonological processing deficit hypothesis. *Reading and Writing, 13*, 31-56.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *The American Psychological Association, 96(1),* 31–42.

Carretti, B., Borella, E., & De Beni, R.D. (2007). Does strategic memory training improve the working memory performance of younger and older adults? *Experimental Psychology, 54*, 311-320.

Castellanos, F.X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Review Neuroscience*, 3, 617-628.

Charach, A., Ickowicz, A., & Schachar, R. (2004). Stimulant treatment over five years: Adherence, effectiveness, and adverse effects. *Journal of the American Academy of Child and Adolescent Psychiatry, 43:5,* 559-567.

Cicerone, K. D., Dahlberg, C., Kalmar, K., Langebahn, D. M., Malec, J. F. & Bergquist, T. F. et al. (2000). Evidence-based cognitive rehabilitation: Recommendations for clinical practice. *Archives of Physical Medicine and Rehabilitation, 81*, 1596-1615.

Conners, C.K., Sitarenios G., Parker J.D., & Epstein, J.N. (1998a). The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, 26, 257-268.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). Routledge

Conners, C.K., Sitarenios G., Parker J.D., & Epstein, J.N. (1998b). Revision and restandardization of the Conners' Teaching Rating Scale (CTRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, 26, 279-291

Conway, A. R. A. & Getz, S. J. (2010). Cognitive Ability: Does working memory training enhance intelligence? *Current Biology, 20,* R362-R364. DOI: 10.1016/j.cub.2010.03.001.

Cornoldi, C., Barbieri, A., Gaiani, C., & Zocchi, S. (1999). Strategic memory deficits in attention deficit disorder with hyperactivity participants: The role of executive processes. *Developmental Neuropsychology*, *15*, 53–71.

Crone, E.A., Wendelken, C., Donohue, S., van Leijenhorst, L., & Bunge, S.A. (2006). Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences, 103*, 9315-9320.

Culbertson, W. C. & Zillmer, E. A. (1999). *The Tower of London, Drexel University, Research Version: Examiner's Manual.* North Tonawanda, NY: Multi-Health Systems.

Dahlin, K. I. E. (2011). Effects of working memory training on reading in children with special needs. *Reading and Writing, 24,* 479-491.

Dahlin, E., Bäckman, L., Neely, A. S. & Nyberg, L. (2009). Training of the executive component of working memory: Subcortical areas mediate transfer effects. *Restorative Neurology and Neuroscience, 27,* 405-419. DOI 10.3233/RNN-2009-0492.

Dahlin, E., Neely, A. S., Larsson, A. & Bäckman, L., (2008). Transfer of learning updating training mediated by the striatum. *Science, 320,* 1510-1512. DOI: 10.1126/science.1155466.

Dahlin, E., Nyberg, L., Bäckman, L. & Neely, A. S. (2008). Executive functioning in younger and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23,* 720-730.

de Fockert, J.W., Rees, G., Frith, C.D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science, 291,* 1803-1806.

de Jong, P. (2006). Understanding normal and impaired reading development: A working memory perspective. In S. Pickering (Ed.), Working memory and education (pp. 33–60). London: Academic Press.

Dempster, F.N. (1981). Memory span: sources of individual and developmental differences. *Psychological Bulletin 89*, 63–100.

Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193-222.

D'Esposito, M., Postle, B. R., Ballard, D. & Lease, J. (1999). Maintenance versus manipulation of information in working memory: An event-related fMRI study. *Brain and Cognition, 41,* 66-86.

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

Dufva, M., Niemi, P., & Voeten, M. J. M. (2001). The role of phonological memory, word recognition, and comprehension skills in reading development: From preschool to grade 2. *Reading and Writing, 14*, 91-117.

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*, 917-928. doi:10.1093/jpepsy/jsp004

Ellis, K. A. & Nathan, P. J. (2001). The pharmacology of human working memory. *International Journal of Neuropsychopharmacology 4,* 299–313.

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

Ericsson, K. A., Krampe, R. T. & Tesch-Römer (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Faraone, S.V., Perlis, R.H., Doyle, A.E., Smoller, J.W., Goralnick, J.J., Holmgren, M.A., et al. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57,* 1313-1323.

Friedman, N. P., Miyake, A., Young, S.E., DeFries, J.C., Corley, R.P. & Hewitt, J.K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychological: General, 137,* 201–225.

Frison, L. & Pocock, S. J. (1992). Repeated measures in clinical trials: Analysis using mean summary statistics and its implications for design. *Statistics in Medicine, 11*, 1685-1704.

Fristad, M.A., Cummins, J., Verducci, J.S., Teare, M. Weller, E.B. & Weller, R.A. (1998). Study IV: concurrent validity of the DSM-IV revised Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology*, *8,* 227.

Fristad, M.A., Glickman, A.R., Verducci, J.S., Teare, M., Weller, E.B., & Weller, R.A. (1998). Study V: Children's interview for psychiatric syndromes (ChIPS): Psychometrics in two community samples. *Journal of Child and Adolescent Psychopharmacology, 8*, 227-245.

Fristad, M.A., Teare, M., Weller, E.B., Weller, R.A., & Salmon, P. (1998). Study III: Development and concurrent validity of the Children's interview for psychiatric syndromes-Parent Version (P-ChIPS): Psychometrics in two community samples. *Journal of Child and Adolescent Psychopharmacology, 8*, 221-226.

Garavan, H., Kelley, D., Rosen, A., Rao S.M. & Stein, E.A. (2000). Practice-related functional activation changes in a working memory task. *Microscopy Research and Technique, 51,* 54-63.

Gathercole, S.E., & Alloway, T.P. (2008). *Working memory and learning: A teacher's guide.* London: Sage Publications.

Gathercole, S.E., Alloway, T.P., Willis, C., & Adams, A. (2006). Working memory in children with reading disabilities. Journal of Experimental Child Psychology, 93, 265–281.

Gathercole, S.E. & Baddeley, A. (1993). *Working Memory and Language.* Hillsdale, NJ: Erlbaum.

Gathercole, S. E., & Pickering, S. J. (2000). Assessment of working memory in six- and seven-year-old children. *Journal of Educational Psychology, 92,* 377–390.

Gau, S.F., Shen, H.Y., Chou, M.C., Tang, C.S., Chiu, Y.N., Gau, C.S. (2006). Determinants of adherence to Methylphenidate and the impact of poor adherence on maternal and family measures. *Journal of Child and Adolescent Psychopharmacology, 16:3,* 286-297.

Gevins, A., Smith, M. E., McEvoy, L. & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex, 7*, 374-385.

Gioia, G., Isquith, P., Guy S.C., & Kenworth, L. (2000a). Behavior rating inventory of executive functions. *Child Neuropsychology, 6*, 235-238.

Gioia, G., Isquith, P., Guy S.C., & Kenworth, L. (2000b). *Behavior Rating of Inventory of Executive Function: Professional Manual. Professional Manual*. Psychological Assessment Resources: Lutz, Florida.

Green, C. T., Long, D. L., Green, D., Iosif, A., Dixon, F. & Miller, M. R. et al. (2012). Will working memory training generalize to improve off-task behavior in children with Attention-Deficit/Hyperactivity Disorder? *Neurotherapeutics, 9*, 639–648.

Gueorguieva, R. & Krystal, J. H. (2004). Progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry, 61*, 310-317.

Hale, T. S., Bookheimer, S., McGough, J. J., Phillips, J. M., McCracken, J. T. (2007). Atypical brain activation during simple & complex levels of processing in adult ADHD: An fMRI study. *Journal of Attention Disorders, 10(4*), 1–16.

Halford, G. S., Cowan, N. & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *TRENDS in Cognitive Science, 11*, 236-242.

Hempel, A., Giesel, F. L., Caraballo, N. M. G., Amann, M., Meyer, H. & Wüstenberg, T. et al. (2004). Plasticity of cortical activation related to working memory during training. *American Journal of Psychiatry, 161*, 745-747.

Hanson, C. A. (2011). *Working memory and short-term memory deficits in individuals with ADHD: A meta-analysis.* Unpublished manuscript.

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

Holmes, J. G., 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, 10,* 1002-1011.

Jacob, C. & Lesch, K. (2006). The Wuerzburg research initiative on adult attention-deficit/hyperactivity disorder (WURIN-AADHD): Multi-layered evaluation of long-term course. *European Archives of Psychiatry and Clinical Neuroscience, 256 (Suppl 1*), I/12–I/20.

Jacobson, N.S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Jaeggi, S. M., Buschkuehl, M., Jonides, J. & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 6829-6833.

Jaus□ovec & Jaus□ovec (2012). Working memory training: Improving intelligence – Changing brain activity. *Brain and Cognition, 79*, 96-106.

Jensen, P. S., Hinshaw, S. P., Kraemer, H. C., Lenora, N., Newcorn, J. H., & Abikoff, H. B. et al. (2001). ADHD comorbidity findings from the MTA study: Comparing comorbid subgroups. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 147–158.

Johnstone, S. J., Roodenrys, S., Phillips, E., Watt, A. J. & Mantz, S. (2010). A pilot study of combined working memory and inhibition training for children with AD/HD. *ADHD Attention Deficit Hyperactivity Disorder, 2,* 31-42.

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.

Kempton, S., Vance, A., Maruff, P., Luk, E., Costin, J., & Pantelis, C. (1999). Executive function and attention deficit hyperactivity disorder: Stimulant medication and better executive function performance in children. *Psychological Medicine 29*, 527–538.

Kendall, P.C., Marrs-Garcia, A., Nath, S.R. & Sheldrick, R.C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology, 67,* 285-299.

Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O., et al. (2006). The prevalence and correlates of adult ADHD in the United States: Results from the National Comorbidity Survey Replication. *American Journal of Psychiatry, 163,* 716–723.

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

Klingberg, T., Fernell, E., Olesen, P.J., Johnson, M., Gustafsson, M.D., Dahlstrom, M.D. et al. (2005). Computerized Training of Working Memory in Children With ADHD--A Randomized, Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 177-186.

Klingberg, T Forssberg, H., & Westerberg, H. (2002a). Increased brain activity in frontal and parietal cortex underlies the development of visuo-spatial working memory capacity during childhood. *Journal of Cognitive Neuroscience, 14,* 1–10.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002b). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24,* 781-791.

Krikorian, R., Bartok, J. & Gay, N. (1994). Tower of London procedures: A standard method and developmental data. *Journal of Clinical and Experimental Neuropsychology, 16,* 840-850.

Kronenberger, W. G., Pisoni, D. B., Henning, S. C., Colson, B. G. & Hazzard, L. M. (2011). Working memory training for children with cochlear implants: A pilot study. *Journal of Speech, Language, and Hearing Research, 54,* 1182-1196.

Kyllonen & Chrystal (1990). Reasoning ability is (little more than) working memory capacity?!. *Intelligence, 14(4),* 389-433.

Levin, H. S., Fletcher, J. M., Kufera, J. A., Harward, H. Lilly, M. A., Mendelsohn, D., et al. (1996). Dimensions of cognition measured by the Tower of London and other cognitive tasks in head-injured children and adolescents. *Developmental Neuropsychology, 12,* 17-34.

Li, D., Sham, P.C., Owen, M.J. & He, L. (2006). Meta-analysis shows significant association between dopamine system genes and attention deficit hyperactivity disorder (ADHD). *Human Molecular Genetics, 15*, 2276-2284.

Llorente, A. M., Voigt, R. G., Berretta, M. C., Rennie, K. A., Fraley, A. K. & Satz, P. et al. (2002). Concurrent validity, factorial validity, and temporal stability of the Children's Color Trails Test 1 & 2 in children with Attention-Deficit/Hyperactivity Disorder (ADHD). Manuscript submitted for publication.

Llorente, A. M., Williams, J., Satz, P. & D'Elia, L. F. (2003). *Children's Color Trails Test: Professional Manual.* Psychological Assessment Resources, Inc: Lutz, FL.

LØhaugen, G. C. C., Antonsen, I., Haberg, A., Gramstad, A., Vik, T. & Brubakk, A., et al. (2011). Computerized working memory training improves function in adolescents born at extremely low birth weight. *The Journal of Pediatrics, 158,* 555-561.e4.

Loosli, S. V., Buschkuehl, M., Perrig, W. J. & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 18:1*, 62-78.

Lo□vdén, M., Bäckman, L., Lindenberger, U., Shaefer, S. & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. Psychological Bulletin, 136, 659-676.

Lunqvist, A., Grundström, K., Samuelsson, K. & Rönnberg, J. (2010). Computerized training of working memory in a group of patients suffering from acquired brain injury. *Brain Injury, 24*, 1173-1183.

Mariani, M.A. & Barkley, R.A.(1997). Neuropsychological and academic functioning in preschools boys with Attention Deficit Hyperactivity Disorder. *Developmental Neuropsychology, 13(1)*, 111-129.

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A Meta-Analysis of Working Memory Impairments in Children With Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 44,* 344-384.

Martinussen, R. & Tannock, R. (2006). Working memory impairments in children with Attention-deficit Hyperactivity Disorder with and without comorbid language disorders. *Journal of Clinical and Experimental Neuropsychology, 28*, 1073-1094.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (in press). Technical Manual. Woodcock-Johnson III Normative Update. Itasca, IL: Riverside Publishing.

McNab, F., Varrone, A., Farde, L., Jucaite, A., Bystritsky, P., Forssberg, H. et al. (2009). Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science, 323,* 800-802.

Mehta, M. A., Goodyear I. M. & Sahakian, B. J. (2004). Methylphenidate improves working memory and set-shifting in AD/HD: Relationships to baseline memory capacity. *Journal of Child Psychology and Psychiatry, 45,* 293-305.

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.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. & Wager, T. D. (2000). The unity and diversity of executive functions and their contribution to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100.

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.

Morton, J. & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen, Developmental Psychopathology (pp. 357–390). New York: John Wiley.

Nigg, J. T. (2006) *What causes ADHD?: Understanding what goes wrong and why*. New York: Guilford Press

Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L. & Rappley, M. D. (2002). Neuropsychological executive functions and DSM-IV ADHD subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry,* 41, 59–66.

Olesen, P.J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience, 7,* 75-79.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S. & Burns, J. S. et al. (2010). Putting brain training to the test. Nature, 465, 775-778.

Owen, A. M., Doyon, J., Petrides, M. & Evans, A. C. (1996). Planning and spatial working memory: A positron emission tomography study in humans. *European Journal of Neuroscience, 8,* 353-364.

Penner, I., Vogt, A., Stöcklin, M., Gschwind, L., Opwis, K. & Calabrese, P. (2012). Computerised working memory training in healthy adults: A comparison of two different training schedules. *Neuropsychological Rehabilitation, iFirst*, 1-18. http://dx.doi.org/10.1080/09602011.2012.686883.

Petersen, S. E., van Mier, H., Fiez, J. A. & Raichle, M. E. (1998). The effect of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Science, 95*, 853-860.

Phillips, L. H., Wynn, V., Gilhooly, K. J., Sala, S. D. & Logie, R. H. (1999). The role of memory in the Tower of London task. *Memory, 7,* 209-231.

Prins, P. J. M., Dovis, S., Ponsioen, A., ten Brink, E. & van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychology, Behavior, and Social Networking, 14,* 115-122.

Raab, G. M., Day, S. & Sales, J. (2000). How to select covariates to include in the analysis of a clinical trial. *Controlled Clinical Trials, 21*, 330-342.

Rapport, M.D., Alderson, R. M., Kofler, M.J., Sarver, D.E., Bolden, J. & Sims, V. (2008). Working memory deficits in boys with Attention-deficit/Hyperactivity Disorder (ADHD): The contribution of central executive and subsystem processes. *Journal of Abnormal Child Psychology, 36,* 825-837.

Rabipour, S. & Raz, A. (2012). Training the brain: Fact and fad in cognitive and behavioral remediation. *Brain and Cognition, 79,* 159-179.

Repovš, G. & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139*, 5-21.

Rosenzweig, M. R. & Bennett, E. L. (1996). Psychobiology of plasticity: Effects of training and experience on brain and behavior. *Behavioural Brain Research, 78,* 57-65.

Roughan, L. & Hadwin, J. A. (2011). The impact of working memory training in young people with social, emotional and behavioural difficulties. *Learning and Individual Differences, 21,* 759–764.

Rutter, M., Bailey, A. & Lord, C. (2003). *The Social Communication Questionnaire Manual.* Western Psychological Services: Los Angeles, CA.

Salminen, T., Strobach, T. & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience, 6, 166,* 1-14. doi: 10.3389/fnhum.2012.00166.

Schmiedek, F., Lovden, 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.

Schneiders, J. A., Opitz, B., Krick, C. M. & Mecklinger, A. (2011). Separating intra-modal and across-modal training effects in visual working memory: An fMRI investigation. *Cerebral Cortex, 21*, 2555-2564.

Schneiders, J. A., Opitz, B., Tang, H., Deng, Y., Xie, C., Li, H. & Mecklinger, A. (2012). The impact of auditory working memory training on the fronto-parietal working memory network. *Frontiers in Human Neuroscience, 6, 173,* 1-14. doi: 10.3389/fnhum.2012.00173.

Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing, 13,* 81–103.

Shaw, P., Eckstrand, K. Sharp, W., Blumenthal, J., Lerch, J.P., Greenstein, D., et al. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences, 104*, 19649-19654.

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

Sieg, K.G., Gaffney, G.G., Preston, D.F. & Hellings, J.A. (1995). SPECT brain imaging abnormalities in attention deficit hyperactivity disorder. *Clinical Nuclear Medicine, 20,* 55-60.

Smith, B.H., Barkley, R.A., & Shapiro, C.J. (2007). Attention-Deficit/Hyperactivity Disorder. In Mash, E.J. & Barkley, R.A. (Eds.), *Assessment of Childhood Disorders* (4th ed. pp. 53-123). New York: The Guilford Press.

Smith, E.E. & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology, 33*, 5-42.

St Clair-Thompson, H. L. (2011). Executive functions and working memory behaviours in children with a poor working memory. *Learning and Individual Differences, 21*, 409-414.

St Clair-Thompson, H., Stevens, R., Hunt, A. & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology, 30,* 203-219. DOI: 10.1080/01443410903509259

Swanson, H. L., Howard, C. B., & Sáez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39(3),* 252–269.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y. & Komuro, N. et al. (2010). Training of working memory impacts structural connectivity. *The Journal of Neuroscience, 30,* 3297-3303.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., & Fukushima, A. et al. (2011). Working memory training using mental calculation impacts regional gray matter of the frontal and parietal regions. *PLoS ONE, 6(8),* e23175. doi:10.1371/journal.pone.0023175.

Tannock, R,. Ickowicz, A. & Schachar, R. (1995). Differential effects of methylphenidate on working memory in ADHD children with or without comorbid anxiety. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 886–896.

Teare, M., Fristad, M.A., Weller, E.B., Weller, R.A., & Salmon, P. (1998a). Study I: Development and criterion validity of the Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology, 8*, 205-211.

Teare, M., Fristad, M.A., Weller, E.B., Weller, R.A., & Salmon, P. (1998b). Study II: Concurrent validity of the DSM-III-R Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of Child and Adolescent Psychopharmacology, 8*, 213-219.

The MTA Cooperative Group (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry, 56, 1073*-1086.

Thorell, L. B., Lindqvist, S. Nutley, S. B., Bohlin, G. & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science, 12,* 106-113.

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

Unsworth, N. & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relations to higher order abilities. Psychological Bulletin, 113, 1038-1066.

Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology, 59,* 920-925. doi: 10.1016/j.jclinepi.2006.02.007.

Verhaeghen, P., Cerella, J. & Basak, C. (2004). Working memory workout: How to expand the focus of attention from one to four items in 10 hours or less. Learning, Memory, and Cognition, 30, 1322–1337. DOI: 10.1037/0278-7393.30.6.1322.

Vogt, A., Kappos, L., Calabrese, P., Stöcklin, M., Gschwind, L. & Opwis, K. et al. (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

Vukovic, R. K. & Siegal, L. (2005). The role of working memory in specific reading comprehension difficulties. In T. P. Alloway & S. E. Gathercole (Eds), *Working memory and neurodevelopmental disorders (*pp. 89-112). New York, NY: Psychology Press.

Wager, T. D. & Smith, E.E. (2003). Neuroimaging studies of working memory: A meta-analysis, *Cognitive Affective & Behavioral Neuroscience, 3,* 255–274.

Wechsler, D. (2003). *The WISC-IV technical and interpretive manual.* San Antonio, TX: Psychological Corporation.

Weller, E.B., Weller, R.A., Rooney, M. & Fristad, M.A. (1999). Children's Interview for psychiatric syndromes: Parent version (P-CHIPS). Washington DC: American Psychiatric Press.

Weller, E.B., Weller, R.A., Rooney, M. & Fristad, M.A. (2000). Children's interview for psychiatric syndromes (ChIPS). *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 76.

Welsh, M. C., Satterlee-Cartmell, T. & Stine, M. (1999). Tower of Hanoi and London: Contribution of working memory and inhibition to performance. *Brain and Cognition, 41,* 231-242.

Westerberg, H., Hirvikoski, T., Forssberg, H., & Klingberg, T. (2004).Visuo-spatial working memory: A sensitive measurement of cognitive deficits in ADHD. *Child Neurology, 10*, 155-161.

Westerberg, H. & Klingberg, T. (2007). Changes in cortical activity after training of working memory-A single-subject analysis. *Physiology & Behavior, 92,* 186-192.

Wiederholt, J. L., & Bryant, B. R. (2001). *GORT 4: Gray Oral Reading Tests Examiner's manual*. Aus- tin, TX: PRO-ED

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V. & Pennington, B. F. (2005). Validity of the executive function theory of attention deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry, 57*, 1336-1346.

Willcutt E. G. & Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: Differences by gender and subtype. *Journal of Learning Disabilities, 33,* 179–191.

Williams, J., Rickert, V., Hogan, J., Zolten, A. J., Satz, P., D'Elia, L. F. et al. (1995). Children's Color Trails. *Archives of Child Neuropsychology, 10*, 211-223.

Witt, M. (2011). School based working memory training: Preliminary finding of improvement in children's mathematical performance. *Advances in Cognitive Psychology, 7,* 7-15. DOI 10.2478/v10053-008-0083-3

Woodcock, R., McGrew, K., Schrank, F, & Mather, N. (2007). Woodcock-Johnson III Normative Update. Rolling Meadows, IL. Riverside Publishing.

Appendix A: Questionnaires

Ohio State University

Child's Name:_____ Birth Date of child : _____ Date Completed: _____

Demographics Questionnaire

Instructions: Please fill out this questionnaire about your child as honestly as possible and to the best of your ability. If you are unsure about how to answer a particular question or you feel uncomfortable responding to a particular question, please skip it.

All information was kept strictly confidential.

---

1. Age: _____ years

2. Sex: ___ Male ___ Female

3. *Please answer **both** 3a **and** 3b.*

(3a) Do you consider your child a person of Mexican, Puerto Rican, Cuban, or South or
Central American culture or origin (regardless of race)?
___ Yes
___ No

(3b) What is your child's racial/ethnic background? Select one or more of the following:
___ American Indian or Native Alaskan
___ Asian
___ Native Hawaiian or Other Pacific Islander
___ Black or African-American
___ Hispanic or Latino
___ White
___ Other (please specify): _____

4. Grade your child is currently in _____

5. Estimated family income per year before taxes:

___ Under $10,000
___ $10,000-$19,999
___ $20,000-29,999
___ $30,000 - $39,999
___ $40,000 - $49,999
___ $50,000 - $59,999
___ $60,000 - $69,999
___ $70,000 - $79,999
___ $80,000 - $89,999
___ $90,000 - $99,999
___ Over $100,000
___ No answer

7. Has your child ever had or been diagnosed with any of the following:

___ Tics or a Tic Disorder
___ Tourette's Disorder
___ Seizures
___ Epilepsy
___ Photo Sensitive Epilepsy

7. Contact Information

Parent Guardian Name(s):_____

_____

Preferred Contact Number:_____

Alternate Contact Number:_____

Email address (optional):_____

Would you prefer evening or weekend scheduling (3 hours):_____

Your Name:_____ Date:_____ Child's Name:_____

Pre DSM-IV Scale

Circle whether your child often does the following and has done it for at least six months.

YES = Your child often does this
NO = Your child does not often do this

1. Often fail to give close attention to details OR makes careless mistakes in school, work, or other activities	YES	NO
2. Often has difficulty sustaining attention in tasks or play activities	YES	NO
3. Often does not seem to listen when spoken to directly	YES	NO
4. Often does not follow through on instructions AND fails to finish school work, chores, or duties in the workplace (NOT due to oppositional behavior or failure to understand instructions	YES	NO
5. Often has difficulty organizing tasks and activities	YES	NO
6. Often avoids, dislikes, OR is reluctant to engage in tasks that required sustained mental effort (Such as schoolwork or homework)	YES	NO
7. Often loses things necessary for tasks or activities (e.g. toys, school assignments, pencils, books, or tools)	YES	NO
8. Is often easily distracted by extraneous stimuli	YES	NO
9. Is often forgetful in daily activities	YES	NO
10. Often fidgets with hands or feet OR squirms in seat	YES	NO
11. Often leaves seat in classroom or in other situations in which remaining seated is expected	YES	NO
12. Often runs about or climbs excessively in situations in which it is inappropriate OR in adolescents, has a feeling of restlessness	YES	NO
13. Often has difficulty playing or engaging in leisure activities quietly	YES	NO

14. Is often “on the go” OR often acts as if “driven by a motor”	YES	NO
15. Often talks excessively	YES	NO
16. Often blurts out answers before questions have been completed	YES	NO
17. Often has difficulty awaiting turn	YES	NO
18. Often interrupts OR intrudes on others (e.g. butts into conversations or games)	YES	NO

Appendix B: Tables

Table 1. Psychiatric Syndromes Assessed by the P-ChIPS

Anxiety Disorders
Acute Stress Disorder
Generalized Anxiety Disorder
Obsessive Compulsive Disorder
Post-traumatic Stress Disorder
Separation Anxiety
Social Phobia
Specific Phobia
Externalizing Disorders
ADHD
Conduct Disorder
Oppositional Defiant Disorder
Mood Disorders
Mania
Hypomania
Major Depressive Disorder
Dysthymic Disorder
Other Disorders
Anorexia
Bulimia
Encopresis
Enuresis
Schizophrenia
Substance Abuse

Table 2. Description of the WM Exercises Included in the Cogmed Working Memory Training Program

Exercise	Description
Rotating Dots	A ring of lights is spinning clockwise. The lights go on in a sequence and participants reproduced the sequence.
Asteroids	Asteroids are moving and rotating in space. They light up in a sequence and participants reproduce the sequence.
Space Whack	Gas comes out of a hole in a sequence. Participants must then put their mouse over the same holes in the same sequence and click on the aliens that come out of the holes.
Grid	Lights are arranged in a four-by-four grid. Participants watch a sequence of lights go on and then reproduce the sequence.
3D Grid	Lights are symmetrically positioned in a 3D "room" with five inner "walls". A sequence of lights goes on and participants reproduce the sequence.
Digit Backwards (visual cues)	Numbers are displayed and a sequence of numbers is vocalized. Participants respond by indicating the numbers in reversed order.
Digit Backwards (no visual cues)	A sequence of numbers is vocalized. Participants respond by indicating the numbers in reversed order.
Grid Rotation	Lights are arranged in a four-by-four grid. Participants watch a sequence of lights go on and then the grid rotates 90° and they reproduce the sequence.
Decoder	Letters are vocalized and a light blinks when each letter is spoken. Then a column of letters is presented below each light and the participants choose the correct letter that was spoken with each light.
Sorter	Doors in a 4 by 4 grid open to reveal numbers and then close. Participants put the numbers in numerical order and click on the doors in that order.
Stabilizer	Letters are vocalized and a light turns on. Then one letter is read and the participant must indicate which light was paired with that letter.
3D Cube	There is a 3-D cube with 5 "walls", all which have 2 panels, except for the back wall, which has 4. The panels light up with different colors in a sequence and the participants must repeat this sequence.

Table 3. Baseline Frequency Comparisons for Experimental and Control Groups

	Experimental Group (*n*)	Control Group (*n*)	Fisher's Exact Test Value*
ADHD Medication			
On medication	9	8	.44
Off medication	8	3	
ADHD Type			
Inattentive	10	8	.69
Combined	7	3	
Presence of Anxiety Disorder			
Yes	5	4	1.00
No	12	7	
Presence of Externalizing Disorder			
Yes	5	8	.05
No	12	3	
Presence of Internalizing Disorder			
Yes	5	4	1.00
No	12	7	
Presence of Mood Disorder			
Yes	1	0	1.00
No	16	11	
Possible Autism Spectrum Disorder			
Yes	2	1	1.00
No	15	9	
Coach			
Coach 1	6	7	.25
Coach 2	11	4	
Presence of Comorbid Disorder			
Yes	8	8	.25
No	9	3	
Ethnicity			
Hispanic	2	1	1.00
Non-Hispanic	15	10	
Race			
Asian American	1	0	.69
Hispanic/Latino	1	1	
White	14	10	
African American	0	0	
Other	1	0	
Sex			
Male	11	9	.42
Female	6	2	
Completer			
Completer	18	11	.14
Non-Completer	7	12	

*when more than 2 groups were compared, the value reported is a chi-squared value

Table 4. Frequencies of Parent-Reported Socioeconomic Statuses

Annual Income	Experimental Group (*n*)	Control Group (*n*)	Chi-Square Value
Under $10,000	0	0	.33
$10,000-$19,999	0	0	
$20,000-29,999	0	0	
$30,000 - $39,999	1	0	
$40,000 - $49,999	0	0	
$50,000 - $59,999	0	0	
$60,000 - $69,999	0	1	
$70,000 - $79,999	2	0	
$80,000 - $89,999	1	3	
$90,000 - $99,999	0	0	
Over $100,000	11	6	
No answer	1	1	

Table 5. Baseline Comparisons between the Experimental and Control Groups

	Experimental Group (n = 17)		Control Group (n = 11)			
Measures	*M*	*SD*	*M*	*SD*	*t*	*p*
Age (months)	148.14	24.08	133.71	32.87	1.34	.19
Total Comorbid Disorders	0.94	1.34	1.55	1.57	-1.09	.29
AWMA Digit Recall	97.00	15.43	90.44	10.97	1.13	.27
AWMA Dot Matrix	93.06	14.45	91.89	18.08	0.18	.86
AWMA Listening Recall	99.88	20.05	98.67	12.64	0.17	.87
AWMA Spatial Recall	98.88	19.43	90.67	10.43	1.18	.25
WISC-IV WM Index	92.47	14.92	87.80	10.51	0.87	.39
BRIEF P WM	65.94	11.13	73.36	6.31	-2.25	.03
BRIEF T WM+	69.13	11.93	73.11	17.93	-0.66	.52
Conners P Hyperactivity/Impulsivity	63.77	14.83	66.82	15.15	-0.53	.60
Conners P Inattention	74.18	11.05	82.55	7.54	-2.20	.04
Conners T Hyperactivity/Impulsivity+	69.40	17.61	61.50	20.33	1.03	.31
Conners T Inattention+	65.40	9.23	66.40	13.66	-0.22	.83
GORT Comprehension	8.57	3.13	9.91	3.33	-1.03	.31
W-J Reading Comprehension	101.82	11.58	95.10	14.92	1.31	.20
Tower Execution Time	93.13	14.80	90.20	17.22	0.46	.65
Tower Move Score	92.63	16.24	102.60	14.21	-1.60	.12
BRIEF P Planning/Organization	62.71	9.85	71.27	9.03	-2.32	.03
BRIEF T Planning/Organization+	67.53	10.11	72.78	13.16	-1.10	.28
BRIEF P GEC	64.00	10.97	70.00	9.99	-1.46	.16
BRIEF T GEC+	73.20	15.91	70.11	18.78	0.43	.67
CCTT Interference Index	1.17	0.87	1.71	1.17	-1.33	.20
Conners P EF	62.71	12.22	72.91	9.75	-2.33	.03

Note. P = Parent, T = Teacher, GEC = Global Executive Control, EF = Executive Functioning

+ experimental group n = 3, control group n = 5

Table 6. Comparison of those who completed post-treatment measures and those who did not on demographic variables

	Completer (*n*)	Non-completer (*n*)	Fisher's Exact Test Value*
ADHD Medication			
On medication	17	15	.22
Off medication	11	4	
ADHD Type			
Inattentive	18	14	.54
Combined	10	5	
Presence of Anxiety Disorder			
Yes	9	9	.37
No	19	10	
Presence of Externalizing Disorder			
Yes	13	13	.56
No	15	11	
Presence of Internalizing Disorder			
Yes	9	9	.37
No	19	10	
Presence of Mood Disorder			
Yes	1	2	.56
No	27	17	
Possible Autism Spectrum Disorder			
Yes	3	3	.67
No	24	15	
Coach			
Coach 1	13	14	.08
Coach 2	15	5	
Presence of Comorbid Disorder			
Yes	16	12	.77
No	12	7	
Ethnicity			
Hispanic	3	1	1.00
Non-Hispanic	25	17	
Race			
Asian American	1	0	.19
Hispanic/Latino	2	1	
White	24	14	
African American	0	3	
Other	1	0	
Sex			
Male	20	15	.74
Female	8	4	
Group			
Experimental	17	7	.14
Control	11	12	

*when more than 2 groups were compared, the value reported is a chi-squared value

Table 7. Frequencies of Parent-Reported Socioeconomic Statuses

Annual Income	Completer (*n*)	Non-completer (*n*)	Chi-Squared Value
Under $10,000	0	0	.11
$10,000-$19,999	0	0	
$20,000-29,999	0	0	
$30,000 - $39,999	1	2	
$40,000 - $49,999	0	0	
$50,000 - $59,999	1	2	
$60,000 - $69,999	1	0	
$70,000 - $79,999	2	1	
$80,000 - $89,999	4	0	
$90,000 - $99,999	0	2	
Over $100,000	17	8	
No answer	2	3	

Table 8. Comparisons between those who did and those who did not complete follow-up measures

	Completers (n = 28)		Non-Completers (n = 19)			
Measures	*M*	*SD*	*M*	*SD*	*t*	*p*
Age (months)	142.47	28.20	152.68	33.51	-1.14	.26
Total Comorbid Disorders	1.18	1.44	2.00	2.05	-1.61	.11
AWMA Digit Recall	94.73	14.18	94.56	19.07	0.03	.97
AWMA Dot Matrix	92.65	15.45	94.88	19.47	-0.41	.68
AWMA Listening Recall	99.46	17.57	103.13	21.39	-0.60	.55
AWMA Spatial Recall	96.04	17.09	95.50	14.83	0.10	.92
WISC-IV WM Index	90.74	13.43	90.35	18.42	0.08	.94
BRIEF P WM	68.86	10.09	74.50	8.87	-1.94	.06
BRIEF T WM	70.63	14.22	73.41	20.22	-0.52	.61
Conners P Hyperactivity/Impulsivity	64.96	14.76	71.11	15.07	-1.39	.17
Conners P Inattention	77.46	10.52	80.37	8.98	-0.98	.33
Conners T Hyperactivity/Impulsivity	66.24	18.75	68.17	16.51	-0.35	.73
Conners T Inattention	65.80	10.95	69.50	13.65	-0.99	.33
GORT Comprehension	9.16	3.22	8.82	2.88	0.35	.73
W-J Reading Comprehension	99.33	13.06	99.53	16.31	-0.05	.97
Tower Execution Time	92.00	15.50	85.68	19.70	1.20	.24
Tower Move Score	96.46	15.98	86.00	16.81	2.12	.04
BRIEF P Planning/Organization	66.07	10.29	69.56	8.19	-1.21	.23
BRIEF T Planning/Organization	69.50	11.36	69.59	18.33	-0.02	.99
BRIEF P GEC	66.36	10.82	70.94	8.95	-1.50	.14
BRIEF T GEC	72.04	16.71	72.12	21.70	-0.01	.99
CCTT Interference Index	1.36	1.00	1.46	0.93	-0.34	.73
Conners P EF	66.71	12.23	71.95	11.67	-1.47	.15

Note. P = Parent, T = Teacher, GEC = Global Executive Control, EF = Executive Functioning

Table 9. Correlations between Pre-Treatment and Post-Treatment Changes on Outcome Measures and Pre-Treatment Scores and Demographic Variables

Treatment Change Variable	Pre-treatment Variable	*r*	*p*
AWMA Dot Matrix	Conners P Learning Problems	-.53	.006
AWMA Dot Matrix	Conners T Learning Problems	-.42	.046
WISC-IV WMI	ADHD Inattentive Symptoms	-.41	.036
BRIEF P WM	BRIEF P WM	-.55*	.002
BRIEF P WM	Conners P Defiance/Aggression	-.64*	<.001
BRIEF P WM	Total Comorbid Diagnoses	-.46	.014
BRIEF T WM	BRIEF T WM	-.84	.009
BRIEF T WM	Conners T Learning Problems	-.84	.010
Conners P Inattention	Conners P Inattention	-.44	.019
Conners P Inattention	Total Comorbid Diagnoses	-.43	.022
Conners P Inattention	Conners P Defiance/Aggression	-.62*	<.001
Conners T Inattention	Conners T Learning Problems	-.77	.026
Conners T Inattention	Total Comorbid Diagnoses	-.72	.045
W-J Oral Language	Total Comorbid Diagnoses	.55*	.003
W-J Reading Comprehension	Conners P Learning Problems	-.47	.016
W-J Reading Comprehension	Conners T Learning Problems	-.47	.020
GORT Comprehension	GORT Comprehension	-.55*	.004
BRIEF P Planning/Organization	BRIEF P Planning/Organization	-.52*	.004
BRIEF P Planning/Organization	Conners P Defiance/Aggression	-.47	.012
BRIEF T Planning/Organization	BRIEF T Planning/Organization	-.91*	.002
BRIEF T Planning/Organization	Conners T Learning Problems	-.90*	.002
TOL Execution Time	TOL Execution Time	-.73*	<.001
TOL Move Score	TOL Move Score	-.48	.014
CCTT Interference Index	CCTT Interference Index	-.78*	<.001
CCTT Interference Index	SCQ	-.46	.020
BRIEF P GEC	BRIEF P GEC	-.64*	<.001
BRIEF P GEC	Conners P Defiance/Aggression	-.68*	<.001
BRIEF T GEC	Conners T Learning Problems	-.92*	.001
BRIEF T GEC	ADHD H/I Symptoms	.75	.031
Conners P EF	Conners P EF	-.48	.010
Conners P EF	Conners P Defiance/Aggression	-.58*	.001
Conners P EF	Total Comorbid Diagnoses	-.38	.043

Note. P = Parent, T = Teacher, EF = executive functioning
*familywise alpha level < .05

Table 10. Descriptive Statistics for Measures of WM

	Experimental Group (n = 17)		Control Group (n = 9)	
	M	*SD*	*M*	*SD*
AWMA Digit Recall				
Pre	97.00	15.43	90.44	10.97
Post	99.47	13.46	94.11	17.40
AWMA Dot Matrix				
Pre	93.06	14.45	91.89	18.08
Post	107.18	23.00	94.33	15.73
AWMA Listening Recall				
Pre	99.88	20.05	98.67	12.64
Post	100.71	20.37	99.67	13.60
AWMA Spatial Recall				
Pre	98.88	19.43	90.67	10.43
Post	104.41	19.14	100.00	8.70
WISC-IV WMI				
Pre	92.47	14.92	87.80	10.51
Post	105.35	16.29	93.40	9.59
WM Composite				
Pre	97.76	13.62	93.92	6.93
Post	103.83	14.44	98.68	7.29
BRIEF P Working Memory				
Pre	65.94	11.13	73.36	6.31
Post	60.59	9.40	65.09	9.86
BRIEF T Working Memory+				
Pre	66.67	19.09	70.80	10.71
Post	78.00	9.64	71.60	4.77

Note. The scores reported for the AWMA are Standard Scores and the scores reported for the BRIEF are T-scores. P = Parent, T = Teacher
+ experimental group n = 3, control group n = 5

Table 11. Post-treatment comparisons between the Experimental and Control Groups on WM

	F	*p*	η_p^2
AWMA Digit Recall			
rANOVA	0.09	.77	<.01
ANCOVA	0.00	.99	<.01
HLM	0.12	.73	
AWMA Dot Matrix			
rANOVA	3.72	.07	.13*
ANCOVA	3.61	.07	.14**
HLM	3.72	.06	
AWMA Listening Recall			
rANOVA	0.00	.98	<.01
ANCOVA	0.00	.98	<.01
HLM	0.06	.81	
AWMA Spatial Recall			
rANOVA	0.94	.34	.04
ANCOVA	0.34	.57	.01
HLM	1.35	.26	
WISC-IV WMI			
rANOVA	4.45	.05	.15**
ANCOVA	5.27	.03	.18**
HLM	2.84	.10	
WM Composite			
rANOVA	0.41	.53	.02
ANCOVA	0.50	.49	
HLM	0.26	.61	
BRIEF P Working Memory			
rANOVA	0.52	.48	.02
ANCOVA	0.16	.69	.01
HLM	0.18	.90	
BRIEF T Working Memory			
rANOVA	1.69	.24	.22**
ANCOVA	3.23	.13	.39**
HLM	2.63	.14	

Note. η_p^2 = partial eta-squared, P = Parent, T = Teacher
†familywise alpha <.05, *moderate effect size, **large effect size

Table 12. Pre-Treatment to Post-Treatment Changes in WM

	Test statistic	*p*	η_p^2
AWMA Digit Recall			
rANOVA	2.28	.14	.09*
Wilcoxon Signed Rank Test	1.64	.10	
HLM	2.80	.11	
AWMA Dot Matrix			
rANOVA	7.48	.01	.24**
Wilcoxon Signed Rank Test	2.77	.01†	
HLM	8.65	.01†	
AWMA Listening Recall			
rANOVA	0.10	.76	<.01
Wilcoxon Signed Rank Test	0.18	.86	
HLM	0.02	.89	
AWMA Spatial Recall			
rANOVA	14.36	<.01†	.37**
Wilcoxon Signed Rank Test	2.96	<.01†	
HLM	14.40	<.01†	
WISC-IV WMI			
rANOVA	28.69	<.01†	.53**
Wilcoxon Signed Rank Test	3.87	<.01†	
HLM	26.58	<.01†	
WM Composite			
rANOVA	28.22	<.01†	.52**
Wilcoxon Signed Rank Test	3.85	<.01†	
HLM	26.11	<.01†	
BRIEF P Working Memory			
rANOVA	11.39	<.01†	.31**
Wilcoxon Signed Rank Test	-2.95	<.01†	
HLM	17.62	<.01†	
BRIEF T Working Memory			
rANOVA	2.25	.19	.27**
Wilcoxon Signed Rank Test	0.94	.35	
HLM	2.06	.19	

Note. η_p^2 = partial eta-squared, P = Parent, T = Teacher
†familywise alpha <.05, *moderate effect size, **large effect size

Table 13. Descriptive Statistics for Measures of ADHD Symptoms

	Experimental Group (n = 17)		Control Group (n = 11)	
	M	*SD*	*M*	*SD*
Conners P Inattention				
Pre	74.18	11.05	82.55	7.54
Post	69.82	11.49	70.64	9.83
Conners T Inattention+				
Pre	59.33	12.42	65.00	12.71
Post	74.00	6.08	68.80	10.06
W-J Oral Language				
Pre	99.00	15.07	94.72	7.80
Post	102.47	17.28	96.64	7.43
Conners P Hyperactivity/Impulsivity				
Pre	63.76	14.83	66.82	15.15
Post	65.53	16.83	64.18	17.77
Conners T Hyperactivity/Impulsivity +				
Pre	74.00	21.93	55.00	18.19
Post	81.33	6.81	54.40	18.24

Note. P = Parent, T = Teacher
+ experimental group n = 3, control group n = 5

Table 14. Post-treatment comparisons between the Experimental and Control Groups on ADHD Symptoms

	F	*p*	η_p^2
Conners P Inattention			
rANOVA	4.69	.04	.15**
ANCOVA	1.86	.19	.07*
HLM	3.24	.08	
Conners T Inattention+			
rANOVA	2.85	.14	.32**
ANCOVA	2.46	.18	.33**
HLM	3.41	.10	
W-J Oral Language			
rANOVA	0.19	.67	.01
ANCOVA	0.35	.56	.01
HLM	0.34	.57	
Conners P Hyperactivity/Impulsivity			
rANOVA	1.79	.19	.06*
ANCOVA	1.70	.21	.06*
HLM	1.25	.27	
Conners T Hyperactivity/Impulsivity			
rANOVA	1.01	.35	.14**
ANCOVA	3.48	.12	.41**
HLM	2.65	.14	

Note. η_p^2 = partial eta-squared, P = Parent, T = Teacher
*moderate effect size, **large effect size

Table 15. Pre-Treatment to Post-Treatment Changes in ADHD Symptoms

	Test statistic	p	η_p^2
Conners P Inattention			
rANOVA	21.71	<.01†	.46**
Wilcoxon Signed Rank Test	-3.25	<.01†	
HLM	28.18	<.01†	
Conners T Inattention+			
rANOVA	8.24	.03	.58**
Wilcoxon Signed Rank Test	1.69	.09	
HLM	7.96	.02	
W-J Oral Language			
rANOVA	2.22	.15	.08*
Wilcoxon Signed Rank Test	1.77	.08	
HLM	2.66	.11	
Conners P Hyperactivity/Impulsivity			
rANOVA	0.07	.79	<.01
Wilcoxon Signed Rank Test	-0.03	.98	
HLM	0.23	.63	
Conners T Hyperactivity/Impulsivity +			
rANOVA	0.73	.43	.11*
Wilcoxon Signed Rank Test	0.14	.89	
HLM	1.04	.34	

Note. η_p^2 = partial eta-squared, P = Parent, T = Teacher
†familywise alpha <.05, *moderate effect size, **large effect size

Table 16. Descriptive Statistics for Measures of Reading Comprehension

	Experimental (n =17)		Control (n = 11)	
	M	*SD*	*M*	*SD*
GORT Comprehension				
Pre	8.57	3.13	9.91	3.33
Post	10.07	2.64	10.45	3.24
W-J Reading Comprehension				
Pre	101.82	11.58	96.89	14.65
Post	104.82	14.35	102.44	20.27

Table 17. Post-treatment comparisons between the Experimental and Control Groups on Reading Comprehension

	F	p	η_p^2
GORT Comprehension			
rANOVA	0.72	.40	.03
ANCOVA	0.11	.74	.01
HLM	0.23	.63	
W-J Reading Comprehension			
rANOVA	0.63	.44	.03
ANCOVA	1.07	.31	.04
HLM	1.04	.32	

Note. η_p^2 = partial eta-squared

Table 18. Comparisons between the Experimental and Control Groups on Reading Comprehension

	F	*p*	η_p^2
GORT Comprehension			
rANOVA	3.33	.08	.13*
Wilcoxon Signed Rank Test	1.88	.06	
HLM	4.27	.05	
W-J Reading Comprehension			
rANOVA	7.01	.01†	.23**
Wilcoxon Signed Rank Test	2.16	.03	
HLM	7.03	.01†	

Note. η_p^2 = partial eta-squared
†familywise alpha <.05, *moderate effect size, **large effect size

Table 19. Descriptive Statistics for Measures of Executive Functions

	Experimental (n = 17)		Control (n = 11)	
	M	*SD*	*M*	*SD*
TOL Move Score				
Pre	92.63	16.24	102.6	14.21
Post	92.50	22.77	99.20	14.05
TOL Execution Time				
Pre	92.67	15.21	90.20	17.22
Post	107.47	10.57	99.00	9.90
BRIEF P Planning/Organization				
Pre	62.71	9.85	71.27	9.03
Post	59.24	10.40	63.82	8.12
BRIEF T Planning/Organization				
Pre	59.67	12.74	71.80	9.47
Post	74.67	4.62	67.00	5.57
CCTT Interference Index				
Pre	1.17	0.87	1.71	1.17
Post	1.40	0.70	1.44	0.66
Conners P EF				
Pre	62.71	12.22	72.91	9.75
Post	60.59	12.89	66.82	9.25
BRIEF P GEC				
Pre	64.00	10.97	70.00	9.99
Post	60.65	9.23	63.55	7.43
BRIEF T GEC+				
Pre	71.00	19.97	65.20	10.21
Post	85.00	17.77	62.80	3.77

Note. P = Parent, T = Teacher, EF = executive functioning, GEC = global executive control

+ experimental group n = 3, control group n = 5

Table 20. Post-treatment comparisons between the Experimental and Control Groups on Executive Functions

	Test Statistic	p	η_p^2
TOL Move Score			
rANOVA	0.14	.72	.01
ANCOVA	0.17	.68	.01
HLM	0.42	.52	
TOL Execution Time			
rANOVA	1.24	.28	.05
ANCOVA	4.56	.04	.17**
HLM	0.12	.74	
BRIEF P Planning/ Organization			
rANOVA	1.18	.29	.04
ANCOVA	0.00	.96	<.01
HLM	0.43	.52	
BRIEF T Planning/Organization			
rANOVA	6.07	.05	.50**
ANCOVA	3.18	.14	.39**
HLM	7.70	.03	
CCTT Interference Index			
rANOVA	1.47	.24	.06*
ANCOVA	0.08	.78	<.01
HLM	1.91	.18	
Conners P Executive Functioning			
rANOVA	0.91	.35	.03
ANCOVA	0.01	.93	<.01
HLM	0.28	.60	
BRIEF P GEC			
rANOVA	0.86	.36	.03
ANCOVA	0.00	.98	<.01
HLM	0.44	.51	
BRIEF T GEC			
rANOVA	7.62	.03	.56**
ANCOVA	13.91	.01	.74**
HLM	11.29	.01†	

Note. η_p^2 = partial eta-squared, P = Parent, T = Teacher, EF = executive functioning, GEC = global executive control

†familywise alpha <.05, *moderate effect size, **large effect size

Table 21. Pre-treatment to post-treatment changes in Executive Functions

	Test Statistic	*p*	*e.s.*
TOL Move Score			
rANOVA	0.16	.69	.01
Wilcoxon Signed Rank Test	-0.31	.75	
HLM	0.38	.54	
TOL Execution Time			
rANOVA	19.22	<.01†	.46**
Wilcoxon Signed Rank Test	3.63	<.01†	
HLM	19.56	<.01†	
BRIEF P Planning/ Organization			
rANOVA	8.88	.01†	.26**
Wilcoxon Signed Rank Test	-2.93	<.01†	
HLM	11.87	<.01†	
BRIEF T Planning/Organization+			
rANOVA	1.61	.25	.21**
Wilcoxon Signed Rank Test	0.56	.58	
HLM	0.76	.41	
CCTT Interference Index			
rANOVA	0.01	.93	<.01
Wilcoxon Signed Rank Test	0.62	.53	
HLM	0.04	,85	
Conners P EF			
rANOVA	3.89	.06	.13*
Wilcoxon Signed Rank Test	-2.02	.04	
HLM	6.15	.02	
BRIEF P GEC			
rANOVA	8.59	.01†	.25**
Wilcoxon Signed Rank Test	-2.52	.01†	
HLM	13.04	<.01†	
BRIEF T GEC+			
rANOVA	3.81	.10	.39**
Wilcoxon Signed Rank Test	0.56	.58	
HLM	3.57	.10	

Note. η_p^2 = partial eta-squared, P = Parent, T = Teacher, EF = executive functioning, GEC = global executive control
†familywise alpha <.05, *moderate effect size, **large effect size

Table 22. Percentage of participants showing clinically significant change and reliable change on each measure where there was statistical significance

	% reaching clinically significant change cutoff post-treatment (pre-treatment)	% showing reliable change
AWMA Dot Matrix	64.3 (52.4)	34.6
Experimental	70.6 (56.5)	47.1
Control	54.5 (47.4)	11.1
AWMA Spatial Recall	71.4 (52.4)	14.8
Experimental	70.6 (69.6)	0.0
Control	72.7 (31.6)	44.4
WISC-IV WMI	66.7 (50.0)	33.3
Experimental	82.4 (60.9)	41.2
Control	40.0 (38.1)	20.0
Working Memory Composite	60.7 (32.1)	17.9
Experimental	58.8 (47.1)	17.7
Control	63.6 (9.1)	18.2
BRIEF P Working Memory	38.1 (19.4)	32.1
Experimental	41.2 (33.3)	17.7
Control	36.4(4.5)	54.6
Conners P Inattention	17.8 (4.2)	28.6
Experimental	17.6 (4.2)	11.8
Control	18.2 (4.3)	54.6
W-J Reading Comprehension	59.2 (52.2)	26.9
Experimental	58.8 (62.5)	23.5
Control	60.0 (40.9)	33.3
TOL Execution Time	74.0 (44.4)	36.0
Experimental	81.2 (47.8)	46.7
Control	63.6(40.9)	20.0
BRIEF P Planning/ Organization	32.2 (25.8)	14.3
Experimental	41.2 (41.2)	5.9
Control	18.2 (9.1)	27.3
BRIEF P GEC	35.7 (19.7)	14.3
Experimental	47.1 (29.4)	17.7
Control	18.2 (9.1)	9.1
BRIEF T GEC	12.5 (24.7)	0.0
Experimental	0.0 (26.7)	0.0
Control	20.0 (22.2)	0.0

Note. P = Parent, T = Teacher, GEC = global executive control