WORKING MEMORY TRAINING IN POST-SECONDARY STUDENTS WITH ATTENTION-DEFICIT/HYPERACTIVITY DISORDER: PILOT STUDY OF THE DIFFERENTIAL EFFECTS OF TRAINING SESSION LENGTH

By

Karizma Mawjee

A thesis submitted in conformity with the requirements for the degree of Master of Arts
Department of Human Development and Applied Psychology
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Karizma Mawjee 2013
Working Memory Training in Post-Secondary Students with Attention-Deficit/Hyperactivity Disorder: Pilot Study of the Differential Effects of Training Session Length

Karizma Mawjee

Master of Arts

Department of Human Development and Applied Psychology
Ontario Institute for Studies in Education
University of Toronto
2013

Abstract

This thesis evaluates the effectiveness of study components in order to aid in design refinements for a larger randomized controlled trial (RCT). A total of 38 post-secondary students with Attention-Deficit/Hyperactivity Disorder (ADHD) were randomized into a waitlist control group, or standard-length (45 minute) or shortened-length (15 minute) WM training group. Criterion measures included the WAIS-IV Digit Span (auditory-verbal WM), CANTAB Spatial Span (visual-spatial WM) and WRAML Finger Windows (visual-spatial WM). Transfer-of-training effects were assessed using indices of everyday cognitive functioning. Participants in the standard- and shortened-length groups showed greater improvements at post-test on auditory-verbal WM and reported fewer cognitive failures in everyday life than waitlist controls. Participants in the standard-length group showed greater improvements in visual-spatial WM at post-test than participants in the other two groups. Preliminary findings suggest that shorter training may have similar beneficial outcomes as documented for the standard-length training, indicating that a larger-scale RCT is warranted.
Acknowledgements

This thesis would not have been possible without the invaluable feedback, guidance and enthusiasm from my supervisor, Dr. Rosemary Tannock. I would also like to thank Dr. Rhonda Martinussen and Dr. Steven Woltering for their helpful advice and suggestions. Many thanks also go out to Howell Gotlieb, Farhad Pantwala and Rachel Goldberg at JVS-Toronto, whose patience and support have been integral to the completion of this study. I am grateful for my research team, Daniel Glizer, Jessica Jung, Brittney Elliot, Victoria Lishak, Alan Rokeach, Samantha Chen, Janice Bertoia, and Palm Soontornsittipong for their tireless help with data entry and collection. Additional thanks to all of the students who agreed to participate and make this study possible. Lastly, I would like to thank my family, friends, and colleagues in the department for their encouragement throughout this process. This thesis research was financially supported by the Canada Research Chairs program (awarded to Tannock) and the Canadian Institutes of Health Research operating Grant (CIHR # 245899, Tannock & Lewis).
# Table of Contents

## Chapter 1: Introduction

1.0 Overview and Rationale for Thesis ................................................................. 1
   1.0.1 My Role in the study.............................................................................. 1
1.1 ADHD ........................................................................................................ 2
1.2 ADHD in Post-secondary Students .............................................................. 3
1.3 Working Memory ......................................................................................... 5
1.4 Working Memory in ADHD ....................................................................... 6
1.5 Working Memory Training ....................................................................... 7
1.6 Working Memory Training in ADHD .......................................................... 9
1.7 Current Study ............................................................................................. 10

## Chapter 2: Research Objectives, Hypothesis, and Methodology

2.0 Pilot Study Purpose .................................................................................... 11
2.1 Objectives and Hypothesis ....................................................................... 11
2.2 Methodology ............................................................................................. 12
   2.2.1 Participants ....................................................................................... 12
   2.2.2 Intervention Program ....................................................................... 15
   2.2.3 Procedure ......................................................................................... 17
   2.2.4 Methods ........................................................................................... 17
   2.2.5 Statistical Analysis .......................................................................... 24

## Chapter 3: Study Manuscript

3.1 Abstract ..................................................................................................... 27
3.2 Introduction ............................................................................................... 28
Chapter 4: Discussion

4.1 Summary ................................................................. 48

4.2 Study Strengths ......................................................... 52

4.3 Limitations ............................................................. 53

4.4 Future Directions ....................................................... 53

4.5 Conclusions ........................................................... 54

References ................................................................. 67
List of Tables

Table 1. Behavioral Studies of Working Memory Training in an ADHD population using an RCT design ................................................................. 56

Table 2.2 Participant Characteristics as Randomized ........................................... 58

Table 2.2 Participant Characteristics as Analyzed ..................................................... 59

Table 3. Descriptive Statistics for Criterion, Near-transfer and Far-transfer measures at pre- and post-test ................................................................. 60

Table 4. Post-treatment Effects of Working Memory Training ................................. 61

Table 5. Confidence Intervals for Post-treatment Effects of Working Memory Training ......................................................................................... 62
List of Figures

Figure 1. Flow of Participants Through the Trial……………………………………………… 63

Figure 2. Post-test Effects of Working Memory Training on Manipulation of
Auditory-Verbal Information ................................................................. 64

Figure 3. Post-test Effects of Working Memory Training on Manipulation of
Visuo-Spatial Information ................................................................. 65

Figure 4. Post-test Effects of Working Memory Training on Everyday Cognitive
Failures ............................................................... 66
Appendix

Appendix A. Description of WM tasks in the Cogmed WM Training Program

........................................................................................................................................... 77
Chapter 1

Introduction

Overview and rationale for thesis

This thesis explores Working Memory (WM) training in post-secondary students with Attention Deficit Hyperactivity Disorder (ADHD). The embedded pilot study was devised to evaluate study components, such as the utility of the proposed treatment and control groups, intake criteria, coach calls and other aspects of the general protocol, so to aid in design refinements for a larger Randomized Controlled Trial (RCT). The first chapter of this thesis reviews the literature on ADHD, working memory, and working memory training. The second chapter summarizes the objectives and hypotheses for the study, and details the methodology employed, which is beyond the scope of what could be included in the publication. Chapter three constitutes the report of the pilot study, written in manuscript format, prepared for publication. Chapter four is a general discussion that expands upon the discussion in the publication manuscript: for instance, it considers the role and effectiveness of WM training from a broader perspective than was possible in the manuscript. The manuscript-based format of this thesis means that there will be some unavoidable overlap of the content of Chapter 3 with that of Chapters 1, 2 and 4. The referenced tables and figures are presented at the end of the end of this thesis.

My role in the study

As the clinical research coordinator for both this pilot study and the ongoing larger scale RCT (Clinicaltrials.gov Identifier: NCT01657721) funded by (Project #: 245899), I was involved in the conceptualization of the study design, creation of study protocol and procedures, creating and maintaining contacts with study collaborators, such as Disability Service Centers at post-
secondary institutions, applying for ethical approval from several post-secondary Research Ethics Boards, participant recruitment and intake, as well as data collection and analysis for this pilot study.

**ADHD**

Attention-Deficit/Hyperactivity Disorder (ADHD) is a prevalent, impairing and heterogeneous neurodevelopmental disorder that is characterized by inattention, hyperactivity and impulsivity (APA, 2000). Based on the fourth edition of the Diagnostics and Statistic Manual (DSM), the following five criteria must be met in order for a diagnosis of ADHD to be established: (1) six or more symptoms of specifically defined criteria for inattention or hyperactivity and impulsivity, (2) symptoms that cause impairment are present prior to age 7, (3) impairment from symptoms are present in at least two settings, (4) clear evidence that the symptoms interfere or reduce quality of social, academic, or occupational functioning, (5) symptoms do not occur exclusively within the course of schizophrenia or another psychiatric disorder, and are not better accounted for by another mental health disorder (APA, DSM-TR-IV, 2000). Additionally, a diagnostic subtype of ADHD may be assigned: Predominantly Inattentive, Predominantly Hyperactive/Impulsive, or Combined subtype. The DSM-5 was published after the completion of this study and includes the following major changes to the diagnostic criteria for ADHD: older adolescents and adults (aged 17 and older) only require the presence of five or more symptoms, age of symptom onset has been increased to 12, and diagnostic specifiers are assigned based on current presentation of symptoms: the subtypes differentiated by DSM-IV turned out to be temporally unstable and to lack validity (Willcutt, Nigg, Pennington, Solanto, Rohde et al., 2012). Also, DSM-5 clusters ADHD under “Neurodevelopmental Disorders” and
not with “Disruptive, Impulse-control and Conduct Disorders”, which conceptually separates ADHD from Oppositional Defiant Disorder and Conduct Disorder for the first time in the DSM.

**ADHD in Post-secondary Students**

ADHD was once thought to be a disorder restricted to childhood, however it is now known that 35-40% of children display symptoms consistent with the disorder into adolescence and adulthood (Biederman, Petty, Evans, Small & Faraone, 2010; Shaw, Malek, Watson, Greenstein, De Rossi et al., 2013). Post-secondary students are an emergent subgroup of the young adult ADHD population (DuPaul, 2009). While it is estimated that 2% to 8% of post-secondary students report clinically significant symptoms of ADHD (DuPaul, 2009), there are several challenges in determining the actual percentage suffering from this disorder. First, there has yet to be a nationally representative epidemiological study conducted to examine the actual incidence of ADHD in post-secondary students (Green & Rabiner, 2013). Second, estimates based on students registered with Post-Secondary Disability Service Centers may be skewed due to (1) underreporting, as students with ADHD are not required to report their diagnosis, or (2) over-reporting, as there are significant incentives for healthy students to feign symptoms; an ADHD diagnosis may entitle students to academic accommodations, such as extra time on tests and access to notes, or allow them to obtain prescriptions for stimulant medication to use for personal gain (Solloman, et al., 2010).

There are also significant barriers in accurately diagnosing ADHD in adulthood. First, clinicians need to establish that ADHD symptoms were present in childhood, which often means relying solely on patients’ retrospective self-reports as it is not always possible to find reliable external sources (Harrison, Edwards & Parker, 2007). Second, it can be difficult to ascertain whether individuals are exaggerating their symptoms or degree of impairment, especially given
the previously mentioned incentives for post-secondary students to feign symptoms (Harrison et al., 2007). These difficulties in diagnosing and determining the incidence of post-secondary students with ADHD have made research with this population challenging. In fact, post-secondary students have mainly received negative attention in the ADHD literature pertaining to symptom feigning and abuse of stimulant medication. Thus, in this study steps were taken during the intake process for this study to validate diagnosis, such as requiring participants be registered with Post-secondary Student Disability Services and asking participants to provide real life examples of impairment and for a significant other to confirm their symptoms.

College students with ADHD are more likely to face social, emotional and academic difficulties than their non-ADHD peers (Blase, Gilbert, Anastopoulos, Costello, Hoyle et al. 2009; Green et al., 2013). The transition to post-secondary education often means the absence of parental support, greater academic demands, new social situations and less structured time which may exacerbate symptoms that were not severe enough to impair functioning in high school. Research has found that compared to their peers, students with ADHD report lower levels of self-esteem, social skills and social adjustment, and greater substance abuse (Baker, Prevatt & Proctor, 2012; Blasé et al., 2009; Shaw-Zirt, Popali-Lehane, Chaplin & Bergman, 2005; Upadhyaya, Rose, Wang, O’Rourke, Sullivan, et al., 2005). Furthermore, they tend to have lower GPA’s, are more likely to face academic probation and are less likely to obtain a college degree or enroll in graduate school (Biederman, Petty, Fried, Kaiser, Dolan, et al., 2008; Frazier, Youngstown, Glutting & Watkins 2007; Heiligenstein, Guenther, Levy, Saviho & Fulwiler, 1999; Mannuzza, Klein, Bessler, Malloy & Mynes, 1997). They also report more difficulties with their concentration, memory, and time management skills than their peers (Kane, Walker, & Schmidt, 2011). These areas of academic difficulties may reflect deficits in executive functioning,
an umbrella term used to define higher-order cognitive processes that are involved in goal-directed behavior that have been theorized to contribute to or exacerbate symptoms of ADHD (Barkley, 1998; Chamberlain, Robbins, Winder-Rhodes, Müller, Sahakian, et al., 2010; Hervey, Epstein & Curry, 2004; Kofler, Rapport, Bolden, Sarver & Raiker, 2010; Martinussen, Hayden, Hogg-Johnson & Tannock, 2005).

**Working Memory**

One domain of Executive Function, Working Memory (WM), has been found to play a particularly important role in sustaining attention, goal attainment and learning (Green, Long, Green, Losif, Dixon, et al., 2012). WM has been defined as the ability to hold and manipulate relevant information in mind for a few seconds in the face of distracting or intrusive thoughts to yield an appropriate outcome (Baddeley, 2003). One of the most prominent theoretical frameworks of WM is Baddeley and Hitch’s (1974) multi-component model, which posits that WM is a limited capacity system that processes information simultaneously across four separable components: (1) the central executive, an attentional control system that filters out irrelevant information and coordinates other WM components, (2) the visuo-spatial sketch pad, a short-term storage system for visual-spatial information, (3) the phonological loop, a short-term storage system for auditory-verbal material, and (4) the episodic buffer, which holds multidimensional chunks of sensory information and allows the other WM components to interface with information from perception and long-term memory.

Individuals vary in the capacity of their WM (Engle, Kane & Tuholski, 1999). This capacity is often measured by complex span tasks that combine memorization of a series of stimuli with concurrent processing tasks, such as listening to a series of numbers and then reciting them in the reverse order. Measures of WM capacity show a strong relationship with
performance in many academic abilities, including reading comprehension, written expression, problem-solving, and mathematical reasoning (Gathercole, Alloway, Willis & Adams, 2006; Swanson & O’Connor, 2009). Moreover, longitudinal research has shown that WM measures are a better predictor of academic success than are IQ scores controlling for initial levels of academic proficiency (Alloway & Alloway, 2010). Thus, deficits in WM may compromise learning outcomes and lead to academic underachievement. WM deficits are common after traumatic brain injury and stroke, and in mental health disorders such as schizophrenia, learning disabilities and ADHD (Gathercole & Alloway, 2006; Martinussen et al., 2005).

**Working Memory in ADHD**

Deficits in Executive Function domains, such as WM, have been implicated in many theoretical models of ADHD. For example, Barkley’s inhibition model (1997) conceptualizes WM deficits in ADHD to stem from inhibitory control impairments, while Rapport’s WM model of ADHD (2001) proposes that WM is a core deficit and that leads to other executive function and behavioral impairments. While WM deficits are not specific to ADHD and that degree of impairment varies within the population (Lambek, Tannock, Dalsgaard, Trillingsgaurd, Damm, et al., 2011; Willcutt, Doyle, Nigg, Faraone & Pennington, 2005), evidence from empirical studies show that individuals with ADHD display moderate to marked impairments in executive function, particularly visuo-spatial and auditory-verbal WM, and that these impairments persist even when other co-morbid disorders or learning disabilities are accounted for (Alloway, Gathercole, Elliott, 2010; Gropper & Tannock, 2009; Martinussen et al., 2005). For post-secondary students with ADHD, executive function impairments can significantly impact their ability to prioritize and organize their homework, memorize facts, complete long-term projects, manage their time and avoid procrastination. Therefore, it is not surprising that WM deficits have
been linked to poor academic outcomes in school-aged children and post-secondary students with ADHD (Gathercole Pickering, Knight & Stegmann, 2004; Gropper & Tannock, 2009). In particular, post-secondary students with ADHD seem to be impaired in auditory-verbal working memory (Gropper & Tannock, 2009).

**Working Memory Training**

Although WM development was once thought to be on a fixed trajectory that was not amenable after early childhood, recent studies have found that WM shows significant neuroplasticity even into adulthood (Klingberg, 2010; Olesen, Westerberg, & Klingberg, 2004). Intensive, systematic and adaptive computerized WM training shows promise for improving WM (Klingberg, Fernell, Olesen, Johnson, Gustafsson et al., 2005). The WM training program used in the present study was developed by CogMed Cognitive Systems (Sweden) to specifically improve WM capacity through daily repetition of visual-spatial and auditory-verbal WM tasks that involve the storage and manipulation of particular sequences of stimuli. One of the main features of this program, which might contribute to training outcomes, is that the difficulty level is automatically adjusted based on trial-by-trial performance to ensure that the program is always taxing the individual’s WM at an optimal level.

WM training has been reported to benefit healthy adults and children, stroke and brain injury patients, as well as children and adolescents with ADHD (Brehmer, Westerberg and Bäckman, 2012; Gibson, Gondoli, Johnson, Steeger, Dobrzenski, et al., 2011; Gray, Chaban, Martinussen, Golberg, Gotlieb, et al., 2012; Green et al., 2012; Lundqvist, Grundstro, Samuelsson & Ronnberg, 2010; Nutley, Soderqvist, Bryde, Humphreys & Klingberg 2011; Westerberg et al., 2007). Studies have found that WM training leads to improvements on non-trained WM tasks (Green et al., 2012; Holmes, Gathercole, & Dunning, 2009; Klingberg,
Forssberg, & Westerberg, 2002; Klingberg et al., 2005; Thorell, Lindqvist, Bergman Nutley, Bohlin, et al., 2009; Westerberg, Jacobaeus, Hirvikoski, Clevberger, Ostensson et al., 2007) (See Table 1). In two of these studies, these beneficial effects lasted three to six months after training was completed (Holmes, Gathercole, & Dunning, 2009; Klingberg et al., 2005). There is also some indication that improvements may generalize to other domains such as attention, reasoning, inhibition and academic skills, such as math. (Green et al., 2012; Johnstone, Roodenrys, Phillips, Watt & Montz, 2010; Holmes et al., 2009; Klingberg et al., 2005).

However, many of these studies have serious methodological limitations, including inconsistent use of valid WM tasks, the use of subjective measures of change, and inclusion of no-contact control groups (Shipstead, Redick, & Engle, 2012). The use of a no-contact control group is particularly problematic, as it does not account for placebo effects; participants are not engaged in the study and may realize that they are not expected to show improvement on post-intervention assessments (Shipstead et al., 2012). To account for this, some WM training studies have employed an active but non-adaptive control group, in which the level of difficulty is fixed rather than being based on performance and the task demands do not exceed keeping more than three items in mind at a time. However, active but non-adaptive training does not provide participants with tangible evidence that their WM abilities are improving as the difficulty level remains unchanged throughout the training process, which may result in similar motivational outcomes as those in no-contact control groups (Shipstead et al., 2012). Therefore, to ascertain that any changes in WM or other cognitive domains are a result of training and not due to placebo effects, an adequate control group must be one that receives some form of adaptive training similar to the treatment groups, to ensure that all participants believe they are receiving
Working Memory Training in ADHD

Studies have found that WM training may be effective for improving WM in children and adolescents with ADHD, but few studies have been conducted to examine the effects of WM training on adults or post-secondary students with ADHD (Gray et al., 2012; Green et al., 2012; Johnstone et al., 2010; Klingberg et al., 2005) (see table 1). One study, by Gropper et al. (accepted August 2013), found that post-secondary students with ADHD who underwent WM training, showed improved performance on several WM tasks, fewer everyday cognitive failures, reduced ADHD symptomatology, and these improvements persisted several months after training. However, this study had several limitations, including the use of a waitlist (i.e., no-contact/inactive) control group, which precluded unequivocal interpretation of the self-reported transfer effects to ADHD symptoms and daily cognitive failures.

In general, there are a limited number of RCTs, particularly ones involving individuals with ADHD, that provide unbiased estimates of training outcomes, or examine whether WM training transfers to other cognitive domains. A recent meta-analysis found that WM training produced short-term training-specific effects that did not generalize to other domains (Melby-Lervåg & Hulme, 2013). Thus, an emerging area of literature now focuses on improvements that can be made to WM training programs to optimize their effects. Alloway et al. (2013) found that training four times a week compared to once a week resulted in significant improvements to WM as well as verbal and nonverbal ability in children with learning disabilities, and that these effects lasted at least 8 months post-training. These findings provide evidence that WM training may be effective but more research is needed to determine the optimum amount of training needing for improvement.
The current study

To date, research on WM training in ADHD has mainly focused on children and adolescents. The success of WM training programs in these populations heavily relies on parental support and motivation. By contrast, adults with ADHD must be intrinsically motivated to undergo WM training and must do so in the face of competing life demands. Post-secondary students in particular, often have to juggle many academic and extra-curricular responsibilities, which can make prioritizing training sessions problematic. This raises the question of whether daily 45-minute training sessions are feasible for students with ADHD; shorter training sessions may be more effective in promoting active engagement and compliance with training activities. No published study to date has investigated whether the length of the daily training sessions influence treatment outcomes. Therefore, the current pilot study evaluates whether shortened daily training sessions (15 minutes) would be an acceptable alternative to the standard length sessions (45 minutes) for a post-secondary ADHD population, whether the outcome measures, intake procedures and overall study design would be appropriate for a larger scale study, and whether a larger scale study using this design is warranted.
Chapter 2

Research Objectives, Hypothesis and Methodology

Pilot Study Purpose

The purpose of this pilot study was to evaluate study components, such as the utility of the proposed treatment and control groups, intake criteria, coach calls and other aspects of the general protocol, so to aid in design refinements for a larger RCT. The data from this pilot study will not be aggregated with those from a larger-scale study. Counselors from Disability Service Centers at several Colleges and Universities were invited to attend a start-up meeting prior to the implementation of the pilot and larger-scale studies. The objective of the start-up meeting was to request their support with recruitment, discuss the feasibility of the study procedures, and seek advice about working with a post-secondary ADHD population. The first two cohorts of participants completed WM training in the summer based on the advice of disability service councilors, as they believed students would have fewer academic commitments during this and training would be more feasible. While we received support from all of the universities and colleges we contacted, the length of processing time for ethical approval differed at each institution, and therefore only university students were included in this pilot study.

Objectives and Hypothesis

In the current study, participants were randomized to receive either the standard-length version (45 minute) of a high intensity, progressive WM training program (CogMed), a shortened-length (15 minute) version of the program, or into a wait-list control group. The inclusion of the shorter-length version not only allowed us to examine ‘dose effects’ of training, but also served as an active control to account for participant expectation and motivation, since the nature of the intervention did not allow for blinding. The major aims of the study were to: 1)
determine the utility of shortened-length WM training by examining the relative compliance, training indices, and effect sizes for training outcomes for the shortened-length training group compared to the standard-length training group; we speculate that shorter training sessions may promote better compliance with the training activities for individuals with ADHD, who are prone to frustration, boredom, and fatigue, 2) evaluate study criteria for validating current ADHD symptomology and detecting possible malingering, 3) explore reasons for any attrition such as inherent differences between compliant and non-compliant participants, possible cohort effects or inadequate intake procedures, and, 4) interpret effect sizes and conduct a power analysis to determine the sample sizes needed for the larger RCT.

**Methodology**

**Participants.** A total of 38 students enrolled in post-secondary education (52.6% male), aged 18-35 ($M=23.39$, $SD=4.02$), with a diagnosis of ADHD participated in this study. A convenience sample was used as we did not specify a sample size apriori, but aimed for about 8-10 participants per group – a sample size that could be feasibly recruited within the first two months of the funding period of the larger-scale RCT. Students were recruited through Information sessions held on campus at 4 universities in a major metropolitan city in Canada. Information sessions were advertised through emails sent from Student Disability Services to students registered with ADHD. In Canada, registration with Student Disability Services requires students to provide comprehensive documentation or undergo a new assessment to confirm their diagnosis. About one third of students (28.9%) were diagnosed with a comorbid learning disability. A total of 23 (60.5%) participants were being treated with medication for their ADHD symptoms. Two were being treated with non-stimulant medication (Atomoxetine), while the majority (n=21) received psychostimulants medication: specifically dextroamphetamine (8),
lisdexamfetamine (4), methylphenidate SR (4) and methylphenidate XR (5). Medication status was not controlled for, experimentally, but participants were advised to maintain their current pharmacological treatment throughout the study. Medication status and dose were recorded at each visit, and three participants reported taking a higher dose of medication at the post-test assessment than at the pre-test assessment.

Semi-structured telephone interviews were conducted to assess students’ eligibility to participate in the study (see figure 1). Inclusion criteria were as follows: 1) previous diagnosis of ADHD, 2) current enrollment in a post-secondary educational institution, 3) registered with Student Accessibility/Disability Services with a diagnosis of ADHD, and 4) between 18 to 35 years of age. Exclusion criteria included: 1) major neurological dysfunction or psychosis, 2) current use of sedating or mood-altering medication other than stimulants provided for ADHD, 3) uncorrected sensory impairment, 4) motor or perceptual handicap that would prevent use of a computer program, or 5) a history of concussion or traumatic brain injury prior to ADHD diagnosis. Exclusion criteria were ascertained from self-report during the intake interview.

To ascertain current clinical symptoms of ADHD, the Adult ADHD Self Report Scale Part A (ASRS-A) was administered orally during the telephone interview and participants were asked to provide real-life examples for each of the six items, to ensure they understood the question and that the reported behavior was a reasonable example of an ADHD symptom. Conventional scoring of the ASRS requires a score of \( \geq 2 \) on items 1-3 and \( \geq 3 \) on items 4-6, with at least four of the six items meeting these criteria to indicate a current symptom profile consistent with a diagnosis of ADHD. We used a number of approaches to validate current ADHD symptomology and detect possible malingering: 1) participants had to be registered with a diagnosis of ADHD at their Post-secondary Disability Service Center, 2) participants needed to
meet the criterion score (detailed above) on the ASRS-A during the telephone interview to be eligible to participate, 3) a significant other completed an adapted version of Adult ADHD Self Report Scale (ASRS) for others (18 items) and the participant completed the full ASRS (18 items) at the pre-test assessment, to show evidence of ongoing elevated symptom count. Significant others, such as a partner, parent, sibling or friend, completed the ASRS for others using a secure, online software program (www.surveymonkey.com). Paired t-tests showed that the differences between ASRS-A (6 items) completed during the telephone interview and the first six items of the full ASRS administered at pretest assessment were not significant, nor were the differences between the adapted ASRS other and the ASRS administered at pretest assessment (all p’s <0.086). Summary data for the self-reported ASRS and other reported ASRS are shown in Table 2.1 and 2.2.

Eligible participants were randomized into one of three treatment arms: standard-length training (45 minutes), shortened-length training (15 minutes) or a delayed-training waitlist control group (to be described in detail below). There were no significant group differences at baseline for age, IQ or severity of ADHD symptoms. Demographic and descriptive characteristics of each treatment arm are shown in Table 2.1 and 2.2. Participants were assessed individually at two different time points; prior to training (T1) and 3 weeks after completing training (T2). On each occasion, participants completed a behavioral assessment, which included a battery of neuropsychological tests and behavioral rating scales, and a neural assessment, where they underwent electroencephalography while completing a battery of executive function tasks. The behavioral and neural assessments were counterbalanced and took a total of five hours at each time point. Data from the neural assessment is reported elsewhere (Woltering, Jung, Liu & Tannock, 2012; Woltering, Liu, Rokeach, & Tannock, 2013) and is not included in this report.
**Intervention Program.** The CogMed WM Training program, developed by CogMed Cognitive Medical Systems AB (Stockholm, Sweden), was used as the intervention program in this study. Thus far, this program has the most empirical evidence to support its efficacy in improving WM (Klingberg, 2010). There are three versions of the CogMed WM training program that all share the same underlying design and algorithms, but have different interfaces to accommodate users based on their age group. The JM version involves 15-minute training sessions and has a carnival themed interface that was designed for pre-school children ages 4-6. The RM version involves 45-minute training sessions for children older than 7, and has a robot themed interface. The QM version involves 45-minute training sessions for adults, and has a basic interface with no theme. Based on an initial survey of participant’s preference, we used the RM version of the Cogmed online-based computer program, as it was perceived to be more engaging than the QM version for adults. The program was implemented through the Cogmed-Pearson’s website (www.cogmed.com) and completed in participants’ homes or place of residence.

The standard RM version of CogMed consists of 12 auditory-verbal and visual-spatial working memory tasks that involve the storage and manipulation of particular sequences of stimuli. An example of a visual-spatial WM task is the ‘3D Grid’, where participants must remember the correct order of boxes that previously lit up on a three dimensional box. An example of an auditory/visual WM task is ‘Hidden’, where a voice recites a series of numbers that participants must reproduce in the reverse order, on a numerical grid (see task descriptions in Appendix A). Positive reinforcement is provided at the end of each trial through computerized verbal feedback. The program involves 25 training sessions to be completed over 5-6 weeks.
Throughout the training process, a certified CogMed coach from a community-based psychological services agency and independent from the research team, conducted weekly telephone-based calls (about 30 min duration) with each participant in the standard-length and shortened-length training group, to provide feedback on their past week’s training performance, advice on training for the following week, address any training challenges, and encourage compliance with the training schedule. Prior to beginning training, participants attended an in-person start-up session, held in a centrally located university conference room, where they were familiarized with their CogMed coaches and the WM training program. Approximately 15 (39.5%) of participants attended a start-up session prior to beginning training.

Participants in the standard-length training group engaged in 45 minutes of training, completing 8 out of the 12 working memory tasks per session that were chosen based on random computer selection of tasks, for a total of 90 WM trials. Those in the shortened-length training group engaged in 15 minutes of training, completing 45 trials of 4 working memory tasks per session, which consisted of two ‘core’ tasks used throughout the training plus two additional tasks that changed during each training session based on random computer selection. Participants were requested to engage in five training sessions per week, taking two days off from training. The standard- and shortened-length training programs both used the adaptive algorithm to adjust task level difficulty, and participants in both groups completed the program in five weeks and received the same amount of coach calls. Thus, the training groups only differed in duration of daily training sessions.

Wait-list control participants did not undergo any training during the 5-week period, but received weekly calls from a certified CogMed coach to control for possible effects of attention and motivation, with each call lasting approximately 30 minutes. Coach calls to waitlist control
participants were a unique component of this study, and one objective of this pilot was to examine whether these types of calls would be feasible in a post-secondary ADHD population. During these calls, wait-list participants were provided with time management, organization, and mnemonic strategies from our ADHD website (www.teachADHD.com). After their T2 assessment, wait-list participants were able to choose between the standard-length and shortened-length working memory training: none opted to undergo training.

**Procedure.** This study was approved by the Institutional Research Ethics Boards of the participating universities, as well as by the community agency providing the WM training program. Informed written consent was obtained from all participants prior to entering the study. Information given to students during recruitment indicated that participants would be randomized into a standard-length (45-minute) training group, shortened-length (15-minute) training group, or wait-list control group. Due to the nature of the treatment arms it was not possible to keep participants or CogMed Coaches blind to condition. However, results of the randomization were not revealed to participants until after the T1 assessment. Participants were run in cohorts of 10-14 to ensure that the CogMed coach was able to provide each participant with adequate individual attention and to better organize the study timeline. Randomization between treatment arms was carried out separately for each cohort. During this study, three cohorts were trained within five months, two cohorts did training during the summer and one cohort did training during the fall semester.

**Measures.**

**Screening Measures.** 1) Vocabulary and Matrix Reasoning subtests from Wechsler’s Abbreviate Scale of Intelligence- Second Edition (WASI-II) were administered at T1 as an estimate of general intellectual ability (Wechsler, 1999). The Vocabulary subtest presents words,
visually and orally, that must be defined, which assesses word knowledge and verbal concept formation. This subtest has 31 items, and each item is scored from 0-2. The Matrix Reasoning subtest has participants complete a matrix or series of patterns, which assesses visual information processing and abstract reasoning skills. This subtest has 31 items and each item is given a score of either 0 (incorrect) or 1 (correct). The Vocabulary and Matrix Reasoning subtest scores are added together and converted into a Partial Scale Intelligence Quotient (PSIQ). The Internal consistency for the WASI PSIQ is 0.93 to 0.98 and test-retest stability is 0.87 to 0.92 (Information for the WASI-II is not yet available; Wechsler, 1999). Participants with a PSIQ below 70 ± 5 were excluded from analyses, as they may have an intellectual disability.

2) Math Fluency subtest from The Woodcock Johnson – Third Edition (WCJ-III; Woodcock, McGrew, Mather & Schrank, 2001), assesses the ability to quickly solve simple mathematical equations. Participants are asked to solve as many math equations as they can in three minutes. The subtest has 160 items, and each item is given a score of either 0 (incorrect) or 1 (correct), which is then converted into a scaled score based on age. Both the Internal consistency and test-retest stability for this battery is excellent with values ranging from .80 to .90 (Woodcock et al., 2001).

3) The Sight Word Reading Efficiency and Phonemic Decoding Efficiency subtests Test of Word Reading Efficiency- Second Edition (TOWRE-II; Wagner, Torgesen & Rashotte 1999) were used as an index of reading difficulties. The Sight Word Reading subtest assesses the number of real words that can be accurately identified in 45 seconds. The subtest has 108 items, and each item is given a score of either 0 (incorrect) or 1 (correct). The Phonemic Decoding Efficiency subtest assesses the number of pronounceable non-real words that can be accurately decoded in 45 seconds. The subtest has 66 items, and each item is given a score of either 0
Raw scores for both subtests are added together to get a Total Word Reading Efficiency Index and then are converted into scaled scores based on age. The Internal consistency for this battery is .86 to .98 and test-retest stability is .82 to .96 (Wagner et al., 1999). Together the Math Fluency Subtest from the WCJ-III and the TOWRE-II were used to screen for learning disabilities: standard scores of 78 or lower on these tests, which represent performance at 7th percentile or lower, may reflect persistent problems in basic academic skills.

4) Symptom Assessment 45 (SA-45) was used to assess general psychiatric symptomology (Maurish, 1999). The questionnaire has nine symptom domain scales: depression, interpersonal sensitivity, hostility, obsessive-compulsive tendencies, psychoticism, paranoid ideation, somatization, and phobic anxiety. There are 45 items that ask participants to what degree certain problems have bothered them in the past 7 days, using 5-point Likert-type scale ranging from not at all (1) to extremely (5). T-scores above 60 indicate elevated psychiatric symptom count.

5) A scale called the ‘GRIT’ to assess perseverance and ambition in terms of long-term goals (Duckworth & Quinn, 2009). The GRIT, which is a colloquial term with no meaning, it is comprised of 18 items that tap ambition, (e.g. I aim to be the best at what I do), perseverance of effort (e.g. I have overcome setbacks to conquer important challenges), and consistency of interest (e.g. I often set goals but later choose to pursue a different one). Total scores range from 18-90. The scale has high internal consistency (.85) and reliability (.79-.80) (Duckworth & Quinn, 2009).

**Outcome Measures.** Outcome measures were categorized into: 1) Compliance Measures: measures that tap compliance with required training intensity, 2) Criterion Measures: assessments that closely resemble tasks from the working memory program, and 3) Near-
Transfer Measures: measures that tap working memory but do not resemble trained tasks, 4) Far-transfer Measures: measures of transfer to everyday functioning, academic performance, and ADHD symptomology. The majority of these measures were selected based on the results of previous studies (Gray et al., 2012; Klingberg et al., 2005; Westerberg et al., 2007). Outcome measures were analyzed using raw scores.

**Compliance Measures.** 1) Duration of training: the number of daily sessions participants completed and number of weeks participants took to complete the 25 required training sessions. 2) Coach calls: the number of the five scheduled coach calls that participants completed. 3) CogMed Training Index Score: this score, which indexes the effort participants put into the training, is calculated by subtracting the results of days 2 and 3 of training from the best two days during the training period; the mean index score for individuals 18 to 65 years is 29 (normal range 15-41), with a higher score signifying good effort put forth during training. 4) Cohort effects: examining whether there was differential attrition between cohorts.

**Criterion Measures.** 1) Digit Span Forwards (DSF), Digit Span Backwards (DSB) and Digit Span Sequencing (DSS) from the Wechsler Adult Intelligence Scale - Fourth Addition (WAIS-IV), were used to assess auditory-verbal working and short-term memory (Wechsler, 2008). The Internal consistency for this battery for ADHD populations is .93 and the internal consistency is .82 (Wechsler, 2008). In these subtests, participants listen to increasingly longer lists of digits, read aloud by the examiner at a rate of one digit per second. During forward trials, the participant must repeat the numbers in the same order that the examiner read them. During the backwards trials, the participant must repeat numbers in the reverse order. During sequencing trials, participants must reorder and repeat the numbers in ascending numerical order.
2) The Spatial Span Forwards (SSF) and Spatial Span Backwards (SSB) from The Cambridge Neuropsychological Testing Automated Battery (CANTAB) was used to assess visual-spatial short-term and working memory (Fray & Sahakian, 1996). The CANTAB battery has high internal consistency (.73-.95) and stability (.60-.70) (Luciana, 2003). CANTAB tasks were administered on a computer using a touch screen monitor. The Spatial Span task presents a set of white squares on the computer screen that change colour in a variable sequence. The participant must then touch the boxes in the same order that they changed colour for the forward subtest and in the reverse order for the backwards subtest. As task demands intensify, the number of boxes is increased from two to nine. However, if the participant makes an error, the next trial remains at the same difficulty level. Spatial span raw scores range from 0-9, with the score representing the highest level at which the participant reproduces at least one correct sequence.

3) The Finger Windows Forwards (FWF) and Finger Windows Backwards (FWB) subtest from The Wide Range Assessment of Memory and Learning (WRAML), was used as another measure of visual-spatial short-term and working memory. The task is administered using an 8 X 11-inch plastic template containing nine asymmetrically located holes, and a pencil. The examiner models a given sequence by placing the end of their pencil into the given holes. Participants imitate the sequence by placing their finger through the holes in the same order for the forwards trail subtest and in the reverse order for the backwards subtest. The total number of correct sequences constitutes the total score, which is then converted into an age-adjusted scaled score. The Internal consistency for this subtest is .81 and test-retest stability is .68 (Adams, 2003).

Near Transfer Measures. 1) The CANTAB Spatial WM task was used to assess strategy skills and visual-spatial working memory (Fray & Sahakian, 1996). This task presents a set of white boxes with blue token hidden underneath them. Through touching the squares and using a
process of elimination, participants have to find one blue token in each box. The participant must not return to a box where a token has already been found. As task demands intensify, the number of boxes is increased from three to eight. Two scores are derived from this task, a between-search errors score and a strategy score. The between-search errors score is computed based on the number of times a participant returns to square where a blue token was previously found, with a higher score representing less errors made. The strategy score evaluates the ability to adopt a consistent search sequence as the number of squares increase, with a low score indicating efficient strategy use, and a high score indicating poor strategy use.

2) The CANTAB Pattern Recognition Memory task assessed visual short-term memory (Fray & Sahakian, 1996). This task presents a series of 12 visual patterns, one at a time, in the centre of the screen. Patterns are not typically geometric shapes and cannot easily be given verbal labels. Participants are required to choose between a pattern they have already seen and a novel pattern. This procedure is then repeated with 12 new patterns for a total of 24 trials. Pattern recognition memory scores range from 0-24.

3) An adapted version of Kahneman’s WM task was designed specially for this study in Microsoft Powerpoint, and was used to assess visual working memory (Kahneman & Peavler, 1969). In this task, subjects mentally add 1 or add 3 to each number in a visually presented sequence of three or four digits ranging from 1 to 9, and say the resulting answer out loud. There are five trials at each of the four levels of difficulty in the following sequence: add 1 to a three digit sequence, add 1 to a four digit sequence, add 3 to a three digit sequence and add 3 to a four digit sequence. The stimulus duration is four seconds and the inter-trial interval is 3 seconds.

**Far Transfer Measures.** 1) The Adult ADHD Self-Report Scale (ASRS) was used to evaluate current manifestation of ADHD symptoms; this scale is based on the DSM-IV and has
high internal consistency and concurrent validity (Adler, Spencer, Faraone, Kessler, Howes et al., 2006). The scale examines symptoms of both hyperactivity (e.g., How often do you fidget or squirm with your hands or feet when you have to sit down for a long time?) and inattention (e.g., How often do you have difficulty concentrating on what people say to you, even when they are speaking to you directly?). The 18 items focus on how often symptoms occur using a 5-point Likert-type scale ranging from never (0) to very often (4), with total scores range from 0-72.

2) The Cognitive Failures Questionnaire (CFQ) was used to assess self-reported errors in memory, perception, and motor function when completing everyday tasks; this questionnaire has been found to have good external validity and stability over time (Broadbent et al., 2011). It is comprised of 35 items that probe for perceptual slips (e.g., Do you fail to notice signposts on the road?), memory slips (e.g., Do you read something and find you haven’t been thinking about it and have to read it again?), and motor slips (e.g. Do you bump into people?). Participants assess the frequency of each cognitive failure over the past 6 months using a 5-point Likert-type scale, with total scores range from 0-75.

3) The Barkley Deficits in Executive Functioning Scale Short Form (BDEFS-SF), was used to evaluate executive functioning deficits in everyday life activities. This scale has been found to have high internal consistency (.91 to.95), and high test-retest reliability (.84) (Barkley, 2011). BDEFS-S uses a 4-point Likert scale that ranges from never or rarely to very often, and scores range from 20-80. It is comprised of 20 items that examine domains such as self-management and time (e.g. procrastinate or put off doing things until the last minute), self-organization (e.g. have difficulty explaining things in their proper order or sequence), self-restraint (e.g. make impulsive comments to others), self-motivation (e.g. unable to work as well
as others without supervision or frequent instruction), and self-regulation of emotion (e.g. have trouble calming myself down when I am emotionally upset).

**Statistical Analysis.** Data were first checked to ascertain the shape of the score distributions using SPSS v.21. A Winsorizing technique described in Tabachnick and Fidell (2007) was used to minimize the effect of at least two outliers in each of the following variables: CANTAB Pattern Recognition Memory, CANTAB Spatial Working Memory Between Errors Score, WAIS backwards, Woodcock Johnson Math Fluency, Kahneman’s addition task, and ASRS. An “As-Treated” approach was used to test for training effects, as there was missing data at post-test (40% attrition) that precluded an Intent-to-Treat analysis of these data. An analysis of covariance (ANCOVA) was conducted with baseline as a covariate, Group (standard-length, shortened-length, and wait-list control) as a between-subjects factor and post-test scores on target indices as the dependent variables. In this pilot study, alpha was adjusted from the conventional .05 level to .1 to increase statistical power to detect treatment effects - ones that might be expected in the larger-scale study adequately powered to detect these effects using a more stringent alpha level. Partial $\eta^2$ values were used to ascertain effect size. According to Vacha-Haase and Thompson (2004), $\eta^2 = .01$ corresponds to a small effect, $\eta^2 = .10$ corresponds to a medium effect, and $\eta^2 = .25$ represents a large effect. Power analyses were conducted using GPower (Faul, Erdfelder, Buchner & Lang, 2009) to determine the optimal sample size needed for a larger randomized control trial. Primary measures of change used for the analysis were two criterion measures and two measures of transfer effects: WAIS Digit Span Backwards subtest, CANTAB Spatial WM Reverse subtest, CANTAB Spatial Working Memory (between errors score) and the ASRS.
The study methodology was designed to, 1) validate participants’ current ADHD symptomology and detecting possible malingering, 2) ensure that training outcomes were appropriately assessed, and 3) detect reasons for any attrition, such as inherent differences between compliant and non-compliant participants or possible cohort effects.
Chapter 3

Working Memory Training in Post-Secondary Students with Attention-Deficit/Hyperactivity Disorder: Pilot Study of the Effects of Training Session Length

Mawjee, K., Tannock, R.¹,²

1 Department of Human Development and Applied Psychology, Ontario Institute for Studies in Education;
2 Neurosciences and Mental Health Research Program, SickKids Hospital

Acknowledgements

We would like to thank JVS – Toronto for their collaboration on this project. Thank you to Howell Gotlieb, Farhad Pantwala and Rachel Goldberg, whose support and enthusiasm were integral to the completion of this study. Many thanks to the students who participated in this study, and to Daniel Glizer, Jessica Jung, and Soyeon Kim for their help with data collection. Funding for this project was provided by the Canadian Institute for Health Research.

Conflicts of Interest

The author declares the following potential conflicts of interest with respect to the research conducted within this thesis: Dr. Tannock, my thesis Supervisor, has been a consultant for Eli Lilly, Purdue, and Shire, for which she has received honoraria that are donated to The Hospital for Sick Children’s Foundation to support ADHD research; Pearson-Cogmed provided the software licences without cost for this study. Ms. Karizma Mawjee reports no financial relationships with commercial interests.
Abstract

Objectives: To evaluate the utility of a shortened-length working memory training program, and the effectiveness of study components such as intake criteria, coach calls and other aspects of the general protocol, so to aid in design refinements for a larger randomized controlled trial (RCT).

Methods: A total of 38 post-secondary students with a previous diagnosis of Attention-Deficit/Hyperactivity Disorder (ADHD) were randomized to receive either the standard 45 minutes of working memory training for 5 days per week for 5 weeks, 15 minutes of working memory training 5 days a week for 5 weeks, or into a waitlist control group, and were evaluated before and 2-3 weeks after completing training. All three groups received weekly advice via telephone from a trained coach. Current symptoms of ADHD were assessed via the Adult Self-Report Scale Part-A (ASRS) during intake and again at pre-test, as well as via a significant other, who completed an adapted ASRS. Criterion measures, which were similar to the trained tasks, included the WAIS-IV Digit Span (auditory-verbal WM), CANTAB Spatial Span (visual-spatial WM) and WRAML Finger Windows (visual-spatial WM). Transfer-of-training effects were assessed using other measures of WM and indices of everyday cognitive functioning. Results: Measures to validate current ADHD symptoms showed there was evidence of ongoing elevated symptom count as reported by the students and their significant other, indicating that malingering in these participants was unlikely. High attrition was attributed in part to inadequate intake procedures. Participants in both the standard- and shortened-length training groups showed greater improvements at post-test compared to waitlist controls on auditory-verbal WM (WAIS-IV Digit Span) and reported fewer cognitive failures in everyday life. Participants in the standard-length training group showed greater improvements in visual-spatial WM (CANTAB Spatial Span) at post-test than those in the shortened-length training and waitlist control group.
**Conclusions:** Preliminary findings suggest that shorter training sessions may yield some of the same beneficial outcomes as documented for the standard-length training, indicating that a larger-scale RCT of shortened versus standard-length WM training is warranted.
**Introduction**

Attention-Deficit/Hyperactivity Disorder (ADHD) is a prevalent, impairing and heterogeneous neurodevelopmental disorder that is characterized by inattention, hyperactivity and impulsivity (APA, 2000). Once thought to be a disorder restricted to childhood, longitudinal research has revealed that about 30-40% of children display symptoms consistent with ADHD into adolescence and adulthood (Biederman, Petty, Evans, Small & Faraone, 2010; Shaw, Malek, Watson, Greenstein, De Rossi, et al., 2013). Post-secondary students are an emergent subgroup of the adult ADHD population, and it is estimated that 2 to 8% of students report clinically significant symptoms of ADHD (DuPaul, 2009). While it is evident that these individuals achieved a certain level of academic success in the past to pursue post-secondary education, they are still more likely to have lower GPA’s and face greater academic difficulties than their non-disordered college peers (Blase, Gilbert, Anastopoulos, Costello, Hoyle et al. 2009; Frazier, Youngstown, Glutting & Watkins 2007). In particular, they report more difficulties with their concentration, memory, and time management skills (Kane, Walker, & Schmidt, 2011).

These areas of difficulty may reflect deficits in executive function, which have been hypothesized to contribute to or exacerbate symptoms of ADHD (Chamberlain, Robbins, Winder-Rhodes, Müller, Sahakian, et al., 2010; Kofler, Rapport, Bolden, Sarver & Raiker, 2010; Martinussen, Hayden, Hogg-Johnson & Tannock, 2005). Working Memory (WM) In particular is one domain of executive function in which individuals with ADHD consistently display moderate to marked impairment (Alloway, Gathercole & Elliott, 2010; Gropper & Tannock, 2009; Kasper, Alderson & Mudec, 2012). WM is the ability to hold and manipulate relevant information in mind for a few seconds in the face of distracting or intrusive thoughts (Baddeley, 2003). It is a limited capacity system that filters out irrelevant information and thus is essential
for sustaining attention, goal attainment and learning (Green, Long, Green, Losif, Dixon, et al., 2012). WM deficits have been linked to poor academic outcomes in school-aged children as well as post-secondary students with ADHD (Alloway et al., 2010; Gropper & Tannock, 2009).

Although WM was once thought to be a fixed genetic trait, recent studies have shown that visual-spatial and auditory-verbal WM may be improved through intensive, adaptive, and computerized training (Klingberg, Fernell, Olesen, Johnson, Gustafsson et al., 2005). Some studies found that WM training also resulted in improvements to generalized domains such as attention, reasoning, inhibition and mathematical skills for children with ADHD, and that these beneficial effects lasted at least 3 to 6 months after training was completed (Klingberg et al., 2005; Green et al., 2012; Johnstone, Roodenrys, Phillips, Watt & Mantz, 2010; Holmes, Gathercole, & Dunning, 2009). While training has been found to improve WM in children with ADHD, healthy adults and children, adults who have suffered a stroke or brain injury, and little research has been done on the adult ADHD population (Brehmer, Westerberg & Bäckman, 2012; Gray, Chaban, Martinussen, Golberg, Gotlieb, et al., 2012; Green et al., 2012; Gibson, Gondoli, Johnson, Steeger, Dobrzenski, et al., 2011; Lundqvist, Grundstro, Samuelsson & Ronnberg, 2010; Nutley, Soderqvist, Bryde, Humphreys & Klingberg et al., 2011; Westerberg, Jacobaeus, Hirvikoski et al., 2007). In the first study to investigate WM training in college students with ADHD, we found that WM training improved performance on several WM tasks, reduced everyday cognitive failures and ADHD symptomatology, and these improvements persisted several months after training (Gropper et al., accepted August 2013). However, this study had several limitations, including the use of a waitlist (i.e., no-contact/inactive) control group, which precluded unequivocal interpretation of the self-reported beneficial effects on ADHD symptoms and daily cognitive failures.
Many WM training studies have similar limitations and thus this area of literature has recently come under scrutiny due to the inconsistent use of valid WM tasks, inclusion of no-contact control groups and the use of subjective measures of change (Shipstead, Redick, & Engle, 2012). Moreover, a recent meta-analysis found that WM training produced short-term training-specific effects that did not generalize to other domains (Melby-Lervåg & Hulme, 2013). Thus, an emerging area of literature now focuses on improvements that can be made to WM training programs to optimize their effects. Alloway et al. (2013) found that training four times a week versus once a week resulted in significant improvements to WM as well as verbal and nonverbal ability in children with learning disabilities, and that these effects lasted at least 8 months post-training. These findings provide evidence that WM training may be effective but more research is needed to determine the optimum amount of training needed for improvement. No published study to date has investigated whether the length of the daily training sessions influence treatment outcomes. Rather, previous studies have used the session length specified on the Cogmed website for the various versions: Cogmed-JM for preschoolers, 10-15 min per day; Cogmed-RM for children and adolescents, 30-45 min per day; and Cogmed-QM for adults, 30-45 min per day. However, we speculate that shorter training sessions may be more effective for individuals with ADHD, who are prone to frustration, boredom, and fatigue: shorter sessions (e.g., 15 min) might promote better concentration and active engagement with the training activities.

In the current study, participants were randomized to receive either the standard-length version (45 minute) of a high intensity, progressive WM training program (CogMed), a shortened-length (15 minute) version of the program, or into a wait-list control group. The inclusion of the shorter-length version not only allowed us to examine ‘dose effects’ of training,
but also served as an active control to account for participant expectation and motivation, since the nature of the intervention did not allow for blinding. The purpose of this pilot study was to evaluate study components, such as the utility of the proposed treatment and control groups, intake criteria, coach calls and other aspects of the general protocol, so to aid in design refinements for a larger RCT. The major aims were to: 1) determine the utility of shortened-length WM training by examining the relative compliance, training indices, and effect sizes for training outcomes for the shortened-length training group compared to the standard-length training group; we speculate that shorter training sessions may promote better compliance with the training activities for individuals with ADHD, who are prone to frustration, boredom, and fatigue, 2) evaluate study criteria for validating current ADHD symptomology and detecting possible malingering, 3) explore reasons for any attrition such as inherent differences between compliant and non-compliant participants, possible cohort effects or inadequate intake procedures, and, 4) interpret effect sizes and conduct a power analysis to determine the sample sizes needed for the larger RCT.

**Methods**

**Participants**

A total of 38 students enrolled in post-secondary education (52.6% male), aged 18-35 ($M=23.39$, $SD=4.02$), with a diagnosis of ADHD participated in this study. A convenience sample was used as we did not specify a sample size a priori, but aimed for there to be about 8-10 participants per group that could be feasibly recruited within the first two months of the study. Students were recruited through Information sessions held on campus at 4 universities in a major metropolitan city in Canada. Information sessions were advertised through emails sent from Student Disability Services to students registered with ADHD. In Canada, registration with
Student Disability Services requires students to provide comprehensive documentation or undergo a new assessment to confirm their diagnosis. About one third of students (28.9%) were diagnosed with a comorbid learning disability. A total of 23 (60.5%) participants were being treated with medication for their ADHD symptoms. Two were being treated with non-stimulant medication (Atomoxetine), while the majority (n=21) received psychostimulants medication: specifically dextroamphetamine (8), lisdexamfetamine (4), methylphenidate SR (4) and methylphenidate XR (5). Medication status was not controlled for, experimentally, but participants were advised to maintain their current pharmacological treatment throughout the study. Medication status and dose were recorded at each visit, and three participants reported taking a higher dose of medication at the post-test assessment than at the pre-test assessment.

Semi-structured telephone interviews were conducted to assess students’ eligibility to participate in the study (see figure 1). Inclusion criteria were as follows: 1) previous diagnosis of ADHD, 2) current enrollment in a post-secondary educational institution, 3) registered with Student Accessibility/Disability Services with a diagnosis of ADHD, and 4) between 18 to 35 years of age. Exclusion criteria included: 1) major neurological dysfunction or psychosis, 2) current use of sedating or mood-altering medication other than stimulants provided for ADHD, 3) uncorrected sensory impairment, 4) motor or perceptual handicap that would prevent use of a computer program, or 5) a history of concussion or traumatic brain injury prior to ADHD diagnosis. Exclusion criteria were ascertained from self-report during the intake interview.

To ascertain current clinical symptoms of ADHD, the Adult ADHD Self Report Scale Part A (ASRS-A) was administered orally during the telephone interview and participants were asked to provide real-life examples for each of the six items, to ensure they understood the question and that the reported behavior was a reasonable example of an ADHD symptom.
Conventional scoring of the ASRS requires a score of ≥ 2 on items 1-3 and ≥ 3 on items 4-6, with at least four of the six items meeting these criteria to indicate a current symptom profile consistent with a diagnosis of ADHD. We used a number of approaches to validate current ADHD symptomology and detect possible malingering: 1) participants had to be registered with a diagnosis of ADHD at their Post-secondary Disability Service Center, 2) participants needed to meet the criterion score (detailed above) on the ASRS-A during the telephone interview to be eligible to participate, 3) a significant other completed an adapted version of Adult ADHD Self Report Scale (ASRS) for others (18 items) and the participant completed the full ASRS (18 items) at the pre-test assessment, to show evidence of ongoing elevated symptom count. Significant others, such as a partner, parent, sibling or friend, completed the ASRS for others using a secure, online software program (www.surveymonkey.com). Paired t-tests showed that the differences between ASRS-A (6 items) completed during the telephone interview and the first six items of the full ASRS administered at pretest assessment were not significant, nor were the differences between the adapted ASRS other and the ASRS administered at pretest assessment (all p’s <0.086). Summary data for the self-reported ASRS and other reported ASRS are shown in Table 2.1 and 2.2.

Eligible participants were randomized into one of three treatment arms: standard-length training (45 minutes), shortened-length training (15 minutes) or a delayed-training waitlist control group (to be described in detail below). There were no significant group differences at baseline for age, IQ or severity of ADHD symptoms. Demographic and descriptive characteristics of each treatment arm are shown in Table 2.1 and Table 2.2. Participants were assessed individually at two different time points; prior to training (T1) and 3 weeks after completing training (T2). On each occasion, participants completed a behavioral assessment,
which included a battery of neuropsychological tests and behavioral rating scales, and a neural assessment, where they underwent electroencephalography while completing a battery of executive function tasks. The behavioral and neural assessments were counterbalanced and took a total of five hours at each time point. Data from the neural assessment is reported elsewhere (Woltering, Jung, Liu, et al., 2012; Woltering, Liu, Rokeach, et al., 2013) and is not included in this report.

**Intervention Program**

The CogMed WM Training program, developed by CogMed Cognitive Medical Systems AB (Stockholm, Sweden), was used as the intervention program in this study. Thus far, this program has the most empirical evidence to support its efficacy in improving WM (Klingberg, 2010). Based on an initial survey of participant’s preference, we used the RM version of the Cogmed online-based computer program, implemented through the Cogmed-Pearson’s website (www.cogmed.com) and completed in participants’ homes or place of residence. The standard RM version of CogMed consists of 12 auditory-verbal and visual-spatial working memory tasks that involve the storage and manipulation of particular sequences of stimuli. For each task, an adaptive algorithm automatically adjusted the difficulty level based on trial-by-trial performance to ensure individuals are always working at the upper limit of their WM capacity. Positive reinforcement is provided at the end of each trial through computerized verbal feedback. The program involves 25 training sessions to be completed over 5-6 weeks.

Throughout the training process, a certified CogMed coach from a community-based psychological services agency and independent from the research team, conducted weekly telephone-based calls (about 30 min duration) with each participant in the standard-length and shortened-length training group, to provide feedback on their past week’s training performance,
advice on training for the following week, address any training challenges, and encourage compliance with the training schedule. Prior to beginning training, participants attended an in-person start-up session, held in a centrally located university conference room, where they were familiarized with their CogMed coaches and the WM training program. Approximately 15 (39.5%) of participants attended a start-up session prior to beginning training.

Participants in the standard-length training group engaged in 45 minutes of training, completing 8 out of the 12 working memory tasks per session that were chosen based on random computer selection of tasks, for a total of 90 WM trials. Those in the shortened-length training group engaged in 15 minutes of training, completing 45 trials of 4 working memory tasks per session, which consisted of two ‘core’ tasks used throughout the training plus two additional tasks that changed during each training session based on random computer selection (see task descriptions in Appendix A). Participants were requested to engage in five training sessions per week, taking two days off from training. The standard- and shortened-length training programs both used the adaptive algorithm to adjust task level difficulty, and participants in both groups completed the program in five weeks and received to the same amount of coach calls. Thus, the training groups only differed in duration of daily training sessions.

Wait-list control participants did not undergo any training during the 5-week period, but received weekly calls from a certified CogMed coach to control for possible effects of attention and motivation, with each call lasting approximately 30 minutes. Coach calls to waitlist control participants were a unique component of this study, and one objective of this pilot was to examine whether these types of calls would be feasible in a post-secondary ADHD population. During these calls, wait-list participants were provided with time management, organization, and mnemonic strategies from our ADHD website (www.teachADHD.com). After their T2
assessment, wait-list participants were able to choose between the standard-length and shortened-length working memory training: none opted to undergo training.

**Procedure**

This study was approved by the Institutional Research Ethics Boards of the participating universities, as well as by the community agency providing the WM training program. Informed written consent was obtained from all participants prior to entering the study. Information given to students during recruitment indicated that participants would be randomized into a standard-length (45-minute) training group, shortened-length (15-minute) training group, or wait-list control group. Due to the nature of the treatment arms it was not possible to keep participants or CogMed Coaches blind to condition. However, results of the randomization were not revealed to participants until after the T1 assessment. Participants were run in cohorts of 10-14 to ensure that the CogMed coach was able to provide each participant with adequate individual attention and to better organize the study timeline. Randomization between treatment arms was carried out separately for each cohort. During this study, three cohorts were trained within five months, two cohorts did training during the summer and one cohort did training during the fall semester.

**Measures**

**Baseline Measures.** 1) Vocabulary and Matrix Reasoning subtests from Wechsler’s Abbreviate Scale of Intelligence- Second Edition (WASI-II) were administered at T1 as an estimate of general intellectual ability (Wechsler, 1999), 2) Math Fluency subtest from The Woodcock Johnson – Third Edition (WCJ-III; Woodcock, McGrew, Mather & Schrank, 2001) and 3) Test of Word Reading Efficiency- Second Edition (TOWRE-II; Wagner, Torgesen & Rashotte, 1999) were used to screen for learning disabilities; 4) Symptom Assessment 45 (SA-45) was used to assess general psychiatric symptomology (Maruish, 1999), and 5) A scale called
the ‘GRIT’ was used to assess ambition, (e.g. I aim to be the best at what I do), perseverance of effort (e.g. I have overcome setbacks to conquer important challenges), and consistency of interest (e.g. I often set goals but later choose to pursue a different one) in terms of long term goals (Duckworth and Quinn, 2009).

**Outcome Measures.** Outcome measures were categorized into: 1) Compliance Measures: measures that tap compliance with required training intensity, 2) Criterion Measures: assessments that closely resemble tasks from the working memory program, and 3) Near-Transfer Measures: measures that tap working memory but do not resemble trained tasks, 4) Far-transfer Measures: measures of transfer to everyday functioning, academic performance, and ADHD symptomology. The majority of these measures were selected based on the results of previous studies (Gray et al., 2012; Klingberg et al., 2005; Westerberg et al., 2007). Outcome measures were analyzed using raw scores.

**Compliance Measures.** 1) Duration of training: the number of daily sessions participants completed and number of weeks participants took to complete the 25 required training sessions. 2) Coach calls: the number of the five scheduled coach calls that participants completed. 3) CogMed Training Index Score: this score, which indexes the effort participants put into the training, is calculated by subtracting the results of days 2 and 3 of training from the best two days during the training period; the mean index score for individuals 18 to 65 years is 29 (normal range 15-41), with a higher score signifying good effort put forth during training. 4) Cohort effects: examining whether there was differential attrition between cohorts.

**Criterion Measures.** 1) Digit Span Forwards (DSF), Digit Span Backwards (DSB) and Digit Span Sequencing (DSS) from the Wechsler Adult Intelligence Scale - Fourth Addition (WAIS-IV), were used to assess auditory-verbal working and short-term memory (Wechsler,
2) The Spatial Span Forwards (SSF) and Spatial Span Backwards (SSB) from The Cambridge Neuropsychological Testing Automated Battery (CANTAB) was used to assess visual-spatial short-term and working memory (Fray & Sahakian, 1996), 3) The Finger Windows Forwards (FWF) and Finger Windows Backwards (FWB) subtest from The Wide Range Assessment of Memory and Learning (WRAML), was used as another measure of visual-spatial short-term and working memory.

**Near Transfer Measures.** 1) The CANTAB Spatial WM task was used to assess strategy skills and visual-spatial working memory (Fray & Sahakian, 1996), 2) The CANTAB Pattern Recognition Memory task assessed visual short-term memory (Fray & Sahakian, 1996), 3) an adapted version of Kahneman’s WM task was used to assess visual working memory. In the latter task, subjects mentally add 1 or add 3 to each number in a visually presented sequence of three or four digits ranging from 1 to 9, and say the resulting answer out loud. There are five trials at each of the four levels of difficulty in the following sequence: add 1 to a three digit sequence, add 1 to a four digit sequence, add 3 to a three digit sequence and add 3 to a four digit sequence. The stimulus duration is four seconds and the inter-trial interval is 3 seconds.

**Far Transfer Measures.** 1) The Adult ADHD Self-Report Scale (ASRS) was used to evaluate current manifestation of ADHD symptoms; this scale is based on the DSM-IV and has high internal consistency and concurrent validity (Adler, Spencer, Faraone, Kessler, Howes, et al. 2006), 2) The Cognitive Failures Questionnaire (CFQ) was used to assess self-reported errors in memory, perception, and motor function when completing everyday tasks; the questionnaire has been found to have good external validity and stability over time (Broadbent et al., 2011), 3) The Barkley Deficits in Executive Functioning Scale Short Form (BDEFS-SF), was used to evaluate executive functioning deficits in everyday life activities (Barkley, 2011).
**Statistical Analysis.** Data were first checked to ascertain the shape of the score distributions using SPSS v.21. A Winsorizing technique described in Tabachnick and Fidell (2007) was used to minimize the effect of at least two outliers in each of the following variables: CANTAB Pattern Recognition Memory, CANTAB Spatial Working Memory Between Errors Score, WAIS backwards, Woodcock Johnson Math Fluency, Kahneman’s addition task, and ASRS. An “As-Treated” approach was used to test for training effects, as there was missing data at post-test (40% attrition) that precluded an Intent-to-Treat analysis of these data. An analysis of covariance (ANCOVA) was conducted with baseline as a covariate, Group (standard-length, shortened-length, and wait-list control) as a between-subjects factor and post-test scores on target indices as the dependent variables. In this pilot study, alpha was adjusted to .1 to increase statistical power and thereby reduce the likelihood of missing treatment effects, which could be expected in the larger-scale study adequately powered to detect these effects using a more stringent alpha level. Partial $\eta^2$ values were used to obtain a rough estimate of effect sizes that could be expected in a larger-scale RCT. According to Vacha-Haase and Thompson (2004), $\eta^2 = .01$ corresponds to a small effect, $\eta^2 = .10$ corresponds to a medium effect, and $\eta^2 = .25$ represents a large effect. Power analyses were conducted using GPower (Faul, Erdfelder, Buchner & Lang, 2009) to determine the optimal sample size needed for a larger randomized control trial. Primary measures of change used for the analysis were two criterion measures and two measures of transfer effects: WAIS Digit Span Backwards subest, CANTAB Spatial WM Reverse subtest, CANTAB Spatial Working Memory (between errors score) and the ASRS.

**Results**

**Baseline Measures.** Baseline findings emphasize the prevalence of WM difficulties in this sample of post-secondary students with ADHD; the mean scaled score for WASI Digit Span
Subtest was 9.47 with 25.9% of scaled scores below the 16%. However, this sample also had a high mean IQ score (112.26), which signifies that in general they are high functioning individuals. Six (15.8%) participants scored below 78 on the Math Fluency Subtest from the WCJ-III, indicating a possible learning disability. However, none of the 11 students that self-reported a learning disability during the intake interview had a score below 78 on either of these subtests. The mean scaled score (61.94) for the SA-45 indicated that this sample had elevated psychiatric symptomology and 54% of participants scoring above 60.

One-way ANOVAs showed there were no differences between the three treatment groups at baseline on WASI-II, WCJ-III, TOWRE-II, SA-45 or GRIT (all p>.1); this showed that our randomization was effective and that the groups had comparable levels of IQ, math fluency, reading fluency, psychiatric symptomology and perseverance (see table 2.1 for all participants that were randomized and table 2.2 for participants used in the analysis).

**Outcome Measures.**

**Compliance Measures.** The attrition rate for this study was high, with 14 of the 38 participants (36.8%) not completing their assigned treatment condition. In the standard-length group, 44.4% (n=8) of participants completed the training in 5.63 weeks ($SD=0.74$, range: 5-7 weeks), with an average Training Index Score of 30.29 ($SD =8.11$, range: 22-42). Participants who did not complete the standard-length training (66.6%) completed an average of 11.6 training sessions ($SD =5.25$, range: 0-16 sessions). In the shortened-length group, 77.8% (n=6) completed the training in 6.29 weeks ($SD =1.38$, range: 5-9 weeks), with an average Training Index Score of 25.71 ($SD =13.43$, range: all but one were between 15-50). Participants who did not complete the shortened-length version (36.4%) completed zero training sessions prior to leaving the study. The completion rate for the standard-length group was significantly different from the shortened-
length group ($\chi^2(1)=2.7, p=.1$). The Training Index Score did not significantly differ between the two training groups ($t(12)=.77, p=.46$). The majority of participants (94%) who completed the training obtained an Training Index Score within the normal range (i.e., a score of 15 to 41). In the waitlist group, 81.8% (n=9) of participants completed all five coach calls: the remaining 22.2% did not receive any calls as they dropped out after the initial T1 assessment. There were no significant differences in attrition rates between cohorts ($\chi^2(2)=.14, p=.93$). A total of 24 participants completed the study requirements and their data was used in the analysis (see table 2.2).

Post hoc analysis (t-tests) of potential baseline differences between compliers and non-compliers revealed significant differences only on subscales of the SA-45 and the GRIT; non-compliers had higher scores on the Paranoia subscale of the SA-45, $t(34, 1)=2.24, p=.03$, and lower scores on the Ambition ($t(x)=2.06, p=.05$) and Perseverance of Effort ($t(x)=2.23, p=.03$) subscales of the GRIT.

**Criterion WM Measures.** Analyses of covariance (ANCOVA) were conducted on the post-test scores with the relevant baseline score as a covariate, Group (standard-length, shortened-length, and wait-list control) as a between-subjects factor. The 95% confident intervals (CI) are shown in Table 5. The ANCOVA for the WAIS-IV Digit Span Backwards was significant ($F(1,24)=7.45, p=.004$) (see figure 2). The strength of the relationship between the group and dependent variable was large, as assessed by a partial $\eta^2$ of .43. Post hoc tests showed that participants in the standard-length group and in the shortened-length group did significantly better at T2 than the waitlist control group (all $p<.059$), and that there were no significant differences between the standard-length group and the shortened-length group ($p=.12$). No differences between the three groups were found at post-test for the WAIS-IV Digit Span
Forwards and Sequencing subtests, however the observed differences at T2 were of medium ($\eta^2=.11$) and small ($\eta^2=.034$) effect size respectively.

The ANCOVA for the CANTAB Spatial Span Backwards subtest was significant ($F(1,24)=4.60, p=.023$) (see figure 3). The strength of the relationship between the group and dependent variable was large, as assessed by a partial $\eta^2$ of .32. Post hoc tests showed that participants in the standard-length group did significantly better than both the shortened-length group and the waitlist control group (all $p<.036$), but there were no significant differences between the shortened-length and the waitlist control group ($p=.44$). Results for the CANTAB Spatial Span Forwards subtest showed no significant effects, however the observed differences at T2 were of medium effect size ($\eta^2=.10$).

With respect to the Finger Windows Forwards and Backwards subtests, there were no significant differences found between the three groups at post-test. However, the observed group differences were of medium effect size (both $p=.17$).

**Near Transfer WM Measures.** No differences were found between the three groups at post-test for any of the near transfer measures (all $p>.16$). However, medium effect sizes were found for the CANTAB Spatial Working Memory Between Errors Score ($\eta^2=.12$) and Strategy Score ($\eta^2=.17$), and small effect sizes were found for CANTAB Pattern Recognition Memory ($\eta^2=.03$) and Kahneman’s Addition Task. ($\eta^2=.04$).

**Far Transfer Measures.** There were no significant differences found between the three groups at post-test for the ASRS, but observed group differences at T2 were of medium effect size for ASRS ($\eta^2=.15$). The ANCOVA conducted for CFQ was significant ($F(1,24) = 3.46, p=.06$) (see figure 4). The strength of the relationship between the group and dependent variable was large, as assessed by a partial $\eta^2$ of .32. Post-hoc analysis showed that participants in the
standard-length group and in the shortened-length group did significantly better at T2 than the waitlist control group (all \( p < .038 \)), and that there were no significant differences between the standard-length group and the shortened-length group (\( p = .83 \)). With respect to BDEFS, no significant differences found between the three groups at post-test (\( p = .96 \)). Confidence intervals for outcome measures are reported in Table 5.

**Power Analysis.** Power analysis for both the CANTAB Spatial Working Memory Between Errors Score and the ASRS revealed that in order for a small effect size to be detected (80% chance) with alpha set at .05, a sample of 18 participants per group would be required. Therefore, given the small sample size (6-9 per group) for this study, analyses were conducted with reduced power.

**Discussion**

The purpose of this pilot study was to evaluate study components, such as the utility of the proposed treatment and control groups, intake criteria, coach calls and other aspects of the general protocol, so to aid in design refinements for a larger RCT. One major objective was to determine the utility of shortened-length WM training (15 minutes) by examining the relative compliance, training indices, and the effect sizes for training outcomes compared to a standard-length group (45 minutes). With regards to the completion rate, significantly more participants completed the shortened-length training (77.8%) compared to the standard-length training group (44%). Importantly, there was no difference in the Training Index Score between the two groups, and the mean score for both groups fell in the expected range. This latter finding indicates that both groups were similarly motivated during training, a concern that has beleaguered previous studies that used an active but non-adaptive control group. An additional aspect of the study also used to control for the effects of motivation and expectation was coach calls to waitlist control
participants; the majority of these participants (81.8%) attended all of their coach calls, indicating that ‘distance coaching’ is feasible in studies involving a post-secondary ADHD population.

Preliminary evidence showed training related effects on two criterion measures of WM. First, the standard-length and shortened-length training groups made greater gains at post-test compared to the waitlist control group on an auditory-verbal WM task (Digit Span Backwards subtest). Second, there were also significant group differences found at post-test for a visual-spatial WM task (CANTAB Spatial Span Backwards subtest), with the standard-length training group performing better at post-test compared to the shortened-length training and waitlist control groups, which showed little to no change over time. Findings also showed that both the standard- and the shortened-length training groups self-reported significantly fewer cognitive failures than did the wait-list control group. As this study was not designed or powered to test efficacy, it is important to apply caution in interpreting these results.

Preliminary results seem to indicate that shortened-length training may be a viable alternative to standard-length training. Both training groups showed comparable effort and motivation during training (based on Training Index), as well as some improvement on some outcome measures (Auditory-Verbal WM, everyday cognitive failures) and compliance did not significantly differ between groups. There is a substantial difference in training demands between the groups, as standard-length WM training involves 18.75 hours of training time per week while shortened-length WM training involves only 6.25 hours. As shortened-length training is easier to schedule and seems to confer some of the same benefits as the standard-length training, it may be a more efficient approach for post-secondary students. These findings
warrant a more rigorous investigation of shortened-length WM training, because if it is effective, the time saved could be allocated to student’s academic and extra-curricular responsibilities.

This pilot study also sought to evaluate the protocol for validating current ADHD symptomology and detecting possible malingering. We used a number of approaches to do this. Specifically, we required that participants: 1) had to be registered with a diagnosis of ADHD at their Post-secondary Disability Service Center, which meant that they must have documented evidence of an ADHD diagnosis, 2) met the criterion score for the six-item scale (Part A) on the Adult ADHD Self Report Scale based upon the initial telephone interview, which indicated a level of severity of symptoms consistent with a diagnosis of ADHD, and 3) showed elevated scores on the total symptom score on the ASRS (18-item) as reported by a second informant (a significant other) as well as on the 18-item self-report ASRS completed at the pre-test assessment, indicating sustained symptomatology across this period of about 2-4 weeks. Analysis of these ASRS scores revealed no significant differences between the ASRS-A (6 items) completed during the telephone interview and the first six items of the 18-item self-reported ASRS administered at pretest assessment, nor between the total scores (based on 18-items) on the second-informant’s ASRS and the participant’s self-reported ASRS administered at pretest assessment. These results show evidence of ongoing elevated symptom count, which indicates that participants were reporting real symptoms and that malingering in these participants was unlikely.

Another objective of the current study was to explore reasons for any attrition, so to fine-tune procedures to maximize compliance in the larger RCT. While the attrition rate for this study was high, with over one-third (about 37%) of participants not completing their assigned treatment condition, analysis revealed inherent differences between completers and non-
completers. Non-completers were more likely to have elevated symptoms of paranoia (SA-45) and lower ambition and perseverance of effort (GRIT), suggesting that they may possess characteristics that increase the likelihood of attrition. These findings suggest that drop-out may not necessarily be related to the WM training requirements per se. Screening for general psychiatric symptomology and persistence during the intake process may be used to alert interventionists to provide additional support to participants who are anxious about this type of intervention or who tend to be poorly motivated to put forth effort, which in turn might reduce attrition.

Other objectives of this study were to obtain rough estimates of effect sizes and determine the sample size needed for a larger confirmatory study. Medium effect sizes were found for the WAIS Digit Span Forwards subtest, CANTAB Spatial Span Forwards, Finger Windows Forwards and Backwards subtest, CANTAB Spatial Working Memory Strategy Score and Between Errors score, and ASRS. Small effect sizes were also found for WAIS Digit Span Sequencing subtest, CANTAB Pattern Recognition Memory Task, and Kahneman’s Addition task. These findings suggest that significant effects could be found using larger sample sizes. Power analysis conducted for a near-transfer measure (CANTAB Spatial Working Memory Between Errors Score) and a far-transfer measure (ASRS) found that in order for a small effect size to be detected (80% chance) as significant at the 5% level, a sample of at least 18 participants per group would be required in a larger, randomized control study.

**Limitations**

While this study had several strengths such as the use of an active control group and rigorous intake procedures to validate ADHD symptoms, there were also several limitations such as high attrition and that post-secondary students from colleges were not included in this study.
Moreover, this pilot study was not designed to be powered for inferential statistical analysis of the training outcomes. Thus, our preliminary reports of group differences in training outcomes need to be interpreted with great caution, and at most they may indicated the magnitude of training outcomes that could be reasonably expected in a larger RCT.

Conclusions

This pilot study is the first to examine the utility of shortened-training sessions for an adaptive WM training program and we did so in a post-secondary ADHD population. We compared compliance rates and preliminary outcomes of a shortened and standard length WM training program. Both training programs used the same adaptive training algorithm over a 5-week period, received the same coaching support, and participants showed comparable levels of motivation during training. Preliminary findings suggest that shorter training sessions may yield higher compliance and some of the same beneficial outcomes as documented for the standard-length training version. We conclude that a larger-scale RCT of shortened versus standard-length WM training is warranted, but that improvements in intake and start-up procedures are required.
Chapter 4

General Discussion

Summary

The purpose of this pilot study was to evaluate study components, such as the utility of the proposed treatment and control groups, intake criteria, coach calls and other aspects of the general protocol, so to aid in design refinements for a larger RCT. As post-secondary students with ADHD have received negative attention in the literature for symptom feigning and abuse of stimulant medication (Solloman et al., 2010), one of the major objectives of this study was to evaluate the protocol for validating current ADHD symptomology and detecting possible malingering. The protocol for this study required participants to: 1) be registered with a diagnosis of ADHD at their Post-secondary Disability Service Center, which meant that they must have documented evidence of an ADHD diagnosis, 2) meet the criterion score for the six-item scale (Part A) on the Adult ADHD Self Report Scale based upon the initial telephone interview, which indicated a level of severity of symptoms consistent with a diagnosis of ADHD, and 3) show elevated scores on the total symptom score on the ASRS (18-item) as reported by a second informant (a significant other) as well as on the 18-item self-report ASRS completed at the pretest assessment, indicating sustained symptomatology across this period of about 2-4 weeks. Analysis of these ASRS scores revealed no significant differences between the ASRS-A (6 items) completed during the telephone interview and the first six items of the 18-item self-reported ASRS administered at pretest assessment, nor between the total scores (based on 18-items) on the second-informant’s ASRS and the participant’s self-reported ASRS administered at the pretest assessment. These results show evidence of ongoing elevated symptom count, which
indicates that participants were reporting real symptoms and that malingering in these participants was unlikely.

Another major objective was to determine the utility of shortened-length WM training (15 minutes) by examining the relative compliance, training indices, and the effect sizes for training outcomes compared to a standard-length group (45 minutes). With regards to the completion rate, significantly more participants completed the shortened-length training (77.8%) compared to the standard-length training group (44%). There was no difference in the Training Index Score between the two groups, and the mean score for both groups fell in the expected range; this indicates that both groups were similarly motivated during training, a concern that has beleaguered previous studies that used an active but non-adaptive control group (Shipstead et al., 2012). Coach calls provided to waitlist control participants, as well as to the two WM training groups, were used to control for the effects of motivation and expectation. The majority of these participants (81.8%) attended all of their coach calls, indicating that this experimental manipulation may be feasible in a post-secondary ADHD population.

Preliminary evidence showed training related effects on two criterion measures of WM: 1) the standard-length and the shortened-length training groups made greater gains at post-test compared to the waitlist control group on an auditory-verbal WM task (Digit Span Backwards subtest), 2) there were also group differences found at post-test for a visual-spatial WM task (CANTAB Spatial Span Backwards subtest), with the standard-length training group performing better at post-test than the shortened-length training and waitlist control groups, which showed little to no change over time. Findings also showed that both the standard- and the shortened-length training groups self-reported significantly fewer cognitive failures than did the wait-list control group, indicating that training may transfer to improvements in cognitive functioning in
everyday life. It is interesting to note that there were no significant effects at post-test for the BDEFS, as this scale was designed to evaluate executive functioning deficits also from an everyday life standpoint. However, when examining these scales at an item level, the CFQ refers to more specific problems that occur in everyday life (e.g. do you leave unimportant letter unopened for days) while the BDEFS tends to refer to general problems in functioning (e.g. inconsistent in the quality or quantity of my work performance). Indicating that the BDEFS may not be sensitive enough to tap subtle changes in everyday functioning.

Preliminary results seem to indicate that shortened-length training may be a viable alternative to standard-length training. Both training groups showed comparable effort and motivation during training (based on Training Index), as well as some improvement on some outcome measures (Auditory-Verbal WM, everyday cognitive failures) and compliance did not significantly differ between groups. There is a substantial difference in training demands between the groups, as standard-length WM training involves 1125 minutes of total training time while shortened-length WM training involves only 375 minutes in total. As shortened-length training is easier to schedule and seems to confer some of the same benefits as the standard-length training, it may be a more efficient approach for post-secondary students. These findings warrant a more rigorous investigation of shortened-length WM training, because if it is effective, the time saved could be allocated to student’s academic and extra-curricular responsibilities. However, it is important to apply caution in interpreting these results, as this study was not designed or powered to test efficacy.

Another objective of the current study was to explore reasons for any attrition, so to fine-tune procedures to maximize compliance in the larger RCT. While recruitment for this study only took a few weeks and students showed enthusiasm to participate, the attrition rate for this
study was high, with over one-third (about 37%) of participants not completing their assigned
treatment condition. Analysis revealed inherent differences between completers and non-
completers; Non-completers were more likely to have elevated symptoms of paranoia (SA-45)
and lower ambition and perseverance of effort (GRIT), suggesting that they may possess
characteristics that increase the likelihood of attrition. These findings suggest that drop-out may
not necessarily be related to the WM training requirements per se. Screening for general
psychiatric symptomology and persistence during the intake process may be used to alert
interventionists to provide additional support to participants who are anxious about this type of
intervention or who tend to be poorly motivated to put forth effort, which in turn might reduce
attrition. Anecdotal observations during the intake interview and from the coaches also suggest
that participants, who may not have fully understood the study parameters, either because they
did not attend a recruitment session or the start-up session with the coaches, seemed to be more
likely to drop-out. This indicates that it may be beneficial for interventionists to provide
additional clarification about the study such as the time commitment required, timeline for
follow-up assessments and expectations for participants.

Finally, we wanted to determine the sample size needed for a larger confirmatory study.
Small effect sizes were found for CANTAB Pattern Recognition Memory Task and Kahneman’s
Addition task, and ASRS. Medium effect sizes were found for the WAIS Digit Span Forwards
subtest, CANTAB Spatial Span Forwards, Finger Windows Forwards and Backwards subtest,
CANTAB Spatial Working Memory Strategy Score and Between Errors score, and ASRS. Small
effect sizes were also found for WAIS Digit Span Sequencing subtest, CANTAB Pattern
Recognition Memory Task, and Kahneman’s Addition task. These findings suggest that
significant effects could be found using larger sample sizes. Power analysis conducted for a near-
transfer measure (CANTAB Spatial Working Memory Between Errors Score) and a far-transfer measure (ASRS) found that in order for a small effect size to be detected (80% chance) as significant at the 5% level, a sample of at least 18 participants per group would be required in a larger, randomized control study.

Study strengths

One of the major strengths of this pilot study was that several approaches were used to validate ADHD symptomology and detect feigning. As post-secondary students with ADHD have mainly received negative attention in the literature regarding symptom feigning and abuse of psychostimulant medication (Solloman et al., 2010), it was important to ensure that the participants recruited for this study were reporting real symptoms of impairment. Another strength was the use of a shorter-length WM training group, as it not only allowed us to examine ‘dose effects’ of training and account for participant expectation, but we were able to explore an alternative treatment option for post-secondary students with ADHD. Moreover, as participants in the shortened-length training group were engaged the same intervention program as the standard-length training group, rather than an alternate cognitive training program, both groups were exposed to the same stimuli, adaptive algorithm and coaching support allowing for equivocal interpretation of training effects. A novel feature of this study were the weekly coach calls to the waitlist control participants, in addition to weekly coach calls to participants in the two WM training groups, which allowed us to control for the effects of motivation.

Limitations

This study had several limitations such as a high attrition rate that the absence of post-secondary students from colleges in this study. Moreover, this pilot study was not designed to be powered for inferential statistical analysis of the training outcomes. Thus, our preliminary reports
of group differences in training outcomes need to be interpreted with great caution, and at most they may indicated the magnitude of training outcomes that could be reasonably expected in a larger RCT.

**Future directions**

A larger confirmatory RCT followed this pilot study and several changes were made to the study protocol to reduce attrition: 1) more time was spent with each participant during intake to thoroughly explain the time commitment required for the WM training, the randomization process, and study procedures for assessments, 2) participants were given timeline sheets that indicated the time period when they would be undergoing WM training and when they would be contacted for follow-up assessments, 3) start-up calls with CogMed coaches were done individually over the phone, rather than in a in-person group format, to ensure each participant received adequate individual attention, 3) participants were given information sheets about their CogMed coach that included a picture and biography of the coach to promote stronger rapport between each participant and their coach, 4) the compensations schedule was shifted so that participants received the largest monetary compensation for their time during the post-test assessment, rather than receiving the same monetary reimbursement for both assessments, to provide motivation to finish with the study after completing the WM training. With these procedures in place, the first cohort in the larger RCT had only 10% attrition. However, it is important to note that for the first cohort for the larger RCT took place during the school year, a time when students have structure and regularity in their routine, which may have made it easier to schedule time for the training. Thus far, despite the advice given to us during the start-up meeting with Disability Services councilors, participants seem more likely to complete the study when the assessments and the WM training take place during the academic term.
Conclusions

This pilot study is the first to examine the utility of shortened-training sessions for an adaptive WM training program and we did so in a post-secondary ADHD population. We compared compliance rates and preliminary outcomes of a shortened and standard length WM training program. Both training programs used the same adaptive training algorithm over a 5-week period, received the same coaching support, and participants showed comparable levels of motivation during training. Preliminary findings suggest that shorter training sessions may yield greater compliance as well as some of the same beneficial outcomes as documented for the standard-length training version. We conclude that a larger-scale RCT of shortened versus standard-length WM training is warranted.
Table 1

**Behavioral Studies of Working Memory Training in an ADHD population using an RCT design**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Population</th>
<th>Measures</th>
<th>Results &amp; Interpretations</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray, Chaban, Martinussen, Goldberg, Gotlieb, Kronitz, Hockenberry, Tannock, 2012; <em>Journal of Child Psychology and Psychiatry</em></td>
<td>Adolescents with severe LD and comorbid ADHD Total N = 60 Age (in years) = 14.3 (1.2)</td>
<td><strong>Criterion</strong> WISC-IV DS, CANTAB SS <strong>Near-transfer</strong> CANTAB SWM, D2 Test of Attention, WM Rating Scale (Alloway, Gathercole, Kirkwood &amp; Elliot, 2009),</td>
<td>1) WM training group improvement on criterion WM measures compared to Math training group, 2) no training effects on near- or far- transfer measures, 3) those who showed biggest improvement on WM training tasks were rated as less inattentive/hyperactive at home by parents</td>
<td>1) Lack of long-term follow up, 2) participants not characterized by type of LD</td>
</tr>
</tbody>
</table>

**WM training group** N=36

**Math training group** N=24

**Far-Transfer**

The IOWA Conners’ Scale, Wide-Range Achievement Test-4-Progress Monitoring Version, The Strengths and Weakness of ADHD-symptoms and Normal-Behaviour Scale (SWAN)
<table>
<thead>
<tr>
<th>Reference</th>
<th>Study Population</th>
<th>Methodology</th>
<th>Criteria</th>
<th>Near-transfer</th>
<th>Far-transfer</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green, Long, Green, Iosif, Dixon, Miller, Fassbender, Schweitzer, 2012; <em>Neurotherapeutics</em></td>
<td>School-aged children with ADHD</td>
<td>Total N=26 Age (in years) = 9.7 (2.2) WM training group N=12 Placebo control group N=14</td>
<td><strong>Criterion</strong> The WM Index score from the WISC-IV <strong>Near-transfer</strong> The Restricted Academic Setting Task (RAST) <strong>Far-transfer</strong> Conners’ Rating Scale</td>
<td>1) WM training group showed significant reductions in off-task ADHD-associated behavior, 2) WM training group significant improved on the WMI of the WISC-IV, 3) No differences between groups on the Conners’</td>
<td>1) Greater number of participants on medication in the WM training group, 2) Modest sample size</td>
<td></td>
</tr>
<tr>
<td>Gibson, Gondoli, Johnson, Steeger, Dobrzenski, Morrissy, 2011; <em>Child Neuropsychology</em></td>
<td>Adolescents with ADHD</td>
<td>Total N=47 Spatial WM Group N=23 Age (in years)= 2.33 (1.17) Verbal WM Group N=24 Age (in years)=12.87 (1.25)</td>
<td><strong>Criterion</strong> Verbal &amp; spatial immediate free recall tasks <strong>Far-transfer</strong> DuPaul ADHD rating scale</td>
<td>1) No differences between spatial and verbal groups, 2) both groups improved maintenance in primary memory, but not recall from secondary memory, 3) Improvement in teacher and parent rated inattentive symptoms</td>
<td>1) Limited outcome measures, 2) no passive/active control group, 3) Some of the ‘verbal’ condition exercises included a spatial component, 4) Main objective was to evaluate program design not efficacy</td>
<td></td>
</tr>
<tr>
<td>Johnstone, Roodenrys, Phillips, Watt, &amp; Mantz, 2010; <em>ADHD Atten Def Hyp Disord.</em></td>
<td>Children and adolescents with ADHD</td>
<td>Total N = 47 Age (in years)=10.7(1.4) High Intensity WM training group N = 15 Low intensity WM training group N = 14</td>
<td><strong>Near-transfer</strong> EEG/ERP baseline task and Go/No Go, Skin conductance</td>
<td>1) Reduced ADHD symptoms (only ‘significant other ratings’ differed between groups, 2) resting EEG changes, 3) Trends: Go/no go, increased arousal, N2 ERP, improved attention alerting</td>
<td>1) No non-WM training control group</td>
<td></td>
</tr>
<tr>
<td>Beck, Hanson, Puffenberger, Benninger, &amp; Benninger, 2010; <em>Journal of Clinical Child and Adolescent Psychiatry</em></td>
<td>Children and adolescents with ADHD</td>
<td>BRIEF, Conners’ PChIPS (for ADHD diagnosis)</td>
<td>1) Parent ratings showed reduced inattentive symptoms and ADHD index (Conners’), 2) parent rating showed improvements on WM, plan/organize and initiate index on BRIEF, 3) significant differences follow up found on Parent Conners’ and all subscales of parent BRIEF, 4) no significant teacher ratings</td>
<td>1) No active control, 2) parents not blind to condition, 3) follow up analysis all t-tests due to missing data (whereas post analysis were repeated measures time*condition)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>Total N = 51 Age (in years) = 11.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klingberg, Fernell, Olesen, Johnson, 2005; <em>Journal of American Academy of Child &amp; Adolescent psychiatry.</em></td>
<td>Children with ADHD/attention problems</td>
<td><em>Criterion</em> Digit span, Span board</td>
<td>1) Post/follow up improvements on Span Board, Digit Span, Stroop, Raven’s Matrices, 2) post/follow up parent ratings lower on inattention and hyperactivity/impulsivity, 3) lower parent rated ODD symptoms and ADHD index scores at follow up</td>
<td>1) No formal ADHD assessment, 2) excluded children with ODD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>Total N = 44</td>
<td><em>Near-transfer</em> Stroop task, Raven’s Matrices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>WM training group N = 22 Control group N = 22</td>
<td><em>Far-transfer</em> Head movements, Connors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klingberg, Forssberg, Westerberg, 2002; <em>Journal of Clinical and Experimental Neuropsychology</em></td>
<td>Experiment 1: Children with ADHD Age (in years) =11</td>
<td><em>Criterion</em> Span board</td>
<td>1) Improvements on trained tasks, visual spatial WM, span board, Raven’s, Stroop, 2) fewer head movements</td>
<td>2) Some children on medication, 2) no measures of changes in everyday life, 3) No follow up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td>Experiment 2: Healthy Adults Age (in years) = 25</td>
<td><em>Near-transfer</em> Stroop task, Choice reaction time tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td><em>Far-transfer</em> Raven’s Matrices, Head movements, Connors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Measures were categorized into criterion, near-transfer and far-transfer at the discretion of the thesis autho*
Table 2.1

Participant Characteristics as Randomized

<table>
<thead>
<tr>
<th></th>
<th>Standard-length group</th>
<th>Shortened-length group</th>
<th>Waitlist Control group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Males (n)</td>
<td>10</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Females (n)</td>
<td>8</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Age (year)</td>
<td>23.67</td>
<td>5.12</td>
<td>21.40</td>
<td>2.70</td>
</tr>
<tr>
<td>WASI IQ$^a$</td>
<td>122.11</td>
<td>12.62</td>
<td>103.00</td>
<td>7.71</td>
</tr>
<tr>
<td>TOWRE$^a$</td>
<td>107.67</td>
<td>14.19</td>
<td>112.20</td>
<td>13.44</td>
</tr>
<tr>
<td>WCJ Math Fluency$^a$</td>
<td>98.33</td>
<td>12.61</td>
<td>92.80</td>
<td>14.04</td>
</tr>
<tr>
<td>ASRS-A Interview$^b$</td>
<td>18.11</td>
<td>2.84</td>
<td>17.4</td>
<td>2.51</td>
</tr>
<tr>
<td>ASRS T1$^b$</td>
<td>51.33</td>
<td>9.41</td>
<td>50.20</td>
<td>10.26</td>
</tr>
<tr>
<td>ASRS Other$^b$</td>
<td>47.11</td>
<td>7.32</td>
<td>47.80</td>
<td>8.58</td>
</tr>
<tr>
<td>SA-45$^a$</td>
<td>61.07</td>
<td>9.82</td>
<td>64.86</td>
<td>9.23</td>
</tr>
<tr>
<td>GRIT$^b$</td>
<td>35.85</td>
<td>4.08</td>
<td>30.00</td>
<td>4.56</td>
</tr>
<tr>
<td>WAIS Digit Span$^a$</td>
<td>9.50</td>
<td>2.23</td>
<td>9.00</td>
<td>2.69</td>
</tr>
</tbody>
</table>

$^a$ standardized scores

$^b$ raw scores
Table 2.2

Participant Characteristics as Analyzed

<table>
<thead>
<tr>
<th></th>
<th>Standard-length group</th>
<th></th>
<th>Shortened-length group</th>
<th></th>
<th>Waitlist Control group</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Males (n)</td>
<td>5</td>
<td></td>
<td>3</td>
<td></td>
<td>5</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Females (n)</td>
<td>3</td>
<td></td>
<td>3</td>
<td></td>
<td>4</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Age (year)</td>
<td>24.25</td>
<td>4.34</td>
<td>21.40</td>
<td>2.70</td>
<td>22.86</td>
<td>3.18</td>
<td>23.13</td>
</tr>
<tr>
<td>WASI IQ&lt;sup&gt;a&lt;/sup&gt;</td>
<td>116.00</td>
<td>14.72</td>
<td>103.00</td>
<td>7.71</td>
<td>114.14</td>
<td>9.77</td>
<td>111.88</td>
</tr>
<tr>
<td>TOWRE&lt;sup&gt;a&lt;/sup&gt;</td>
<td>106.75</td>
<td>12.84</td>
<td>112.20</td>
<td>13.44</td>
<td>97.43</td>
<td>9.03</td>
<td>104.08</td>
</tr>
<tr>
<td>WCJ Math Fluency&lt;sup&gt;a&lt;/sup&gt;</td>
<td>96.25</td>
<td>5.19</td>
<td>92.80</td>
<td>14.04</td>
<td>90.29</td>
<td>13.90</td>
<td>94.08</td>
</tr>
<tr>
<td>ASRS-A Interview&lt;sup&gt;b&lt;/sup&gt;</td>
<td>18.00</td>
<td>1.83</td>
<td>17.40</td>
<td>2.51</td>
<td>18.00</td>
<td>3.51</td>
<td>17.21</td>
</tr>
<tr>
<td>ASRS T1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>49.00</td>
<td>10.23</td>
<td>50.20</td>
<td>10.26</td>
<td>51.86</td>
<td>11.02</td>
<td>49.61</td>
</tr>
<tr>
<td>ASRS Other&lt;sup&gt;b&lt;/sup&gt;</td>
<td>46.25</td>
<td>5.25</td>
<td>47.80</td>
<td>8.58</td>
<td>43.57</td>
<td>8.22</td>
<td>45.56</td>
</tr>
<tr>
<td>SA-45&lt;sup&gt;a&lt;/sup&gt;</td>
<td>55.29</td>
<td>5.77</td>
<td>66.50</td>
<td>8.92</td>
<td>60.89</td>
<td>9.09</td>
<td>60.64</td>
</tr>
<tr>
<td>GRIT&lt;sup&gt;b&lt;/sup&gt;</td>
<td>37.17</td>
<td>3.31</td>
<td>30.00</td>
<td>4.56</td>
<td>27.67</td>
<td>7.21</td>
<td>30.17</td>
</tr>
<tr>
<td>WAIS Digit Span</td>
<td>9.25</td>
<td>1.83</td>
<td>9.43</td>
<td>2.93</td>
<td>9.00</td>
<td>2.47</td>
<td>9.21</td>
</tr>
</tbody>
</table>

a. standardized scores  
b. raw scores
Table 3

Descriptive Statistics for Criterion, Near-transfer and Far-transfer measures at pre- and post-test

<table>
<thead>
<tr>
<th>Measure</th>
<th>Standard-length Training Group (n=8)</th>
<th>Shortened-length Training Group (n=7)</th>
<th>Wait-list Control Group (n=9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-test M  SD</td>
<td>Post-test M  SD</td>
<td>Pre-test M  SD</td>
</tr>
<tr>
<td>WAIS Digit Span: Forwards</td>
<td>9.38  1.30</td>
<td>12.50  2.67</td>
<td>10.57  1.62</td>
</tr>
<tr>
<td>WAIS Digit Span: Backwards</td>
<td>8.75  2.77</td>
<td>12.75  2.49</td>
<td>7.86  2.27</td>
</tr>
<tr>
<td>WAIS Digit Span: Sequencing</td>
<td>8.75  1.49</td>
<td>10.12  2.10</td>
<td>8.86  2.27</td>
</tr>
<tr>
<td>CANTAB Spatial Span: Forwards</td>
<td>6.75  2.50</td>
<td>7.88  1.13</td>
<td>6.57  1.40</td>
</tr>
<tr>
<td>CANTAB Spatial Span: Backwards</td>
<td>5.88  1.64</td>
<td>7.75  1.28</td>
<td>6.29  1.98</td>
</tr>
<tr>
<td>WRAML Finger Windows: Forwards</td>
<td>18.38  3.25</td>
<td>22.13  2.23</td>
<td>15.29  4.31</td>
</tr>
<tr>
<td>CANTAB Spatial Working Memory: Between Errors Score</td>
<td>26.88  6.45</td>
<td>22.50  4.21</td>
<td>33.86  4.49</td>
</tr>
<tr>
<td>CANTAB Pattern Recognition Memory</td>
<td>78.13  19.76</td>
<td>83.85  15.18</td>
<td>83.93  14.52</td>
</tr>
<tr>
<td>Kahneman’s Addition Task</td>
<td>16.75  4.03</td>
<td>18.25  2.19</td>
<td>16.14  4.88</td>
</tr>
<tr>
<td>ASRS</td>
<td>47.71  7.87</td>
<td>46.57  8.44</td>
<td>49.29  9.69</td>
</tr>
<tr>
<td>CFQ</td>
<td>51.66  5.92</td>
<td>44.67  6.38</td>
<td>58.75  12.95</td>
</tr>
<tr>
<td>BDEFS</td>
<td>41.14  11.64</td>
<td>41.71  12.15</td>
<td>54.71  8.60</td>
</tr>
</tbody>
</table>

Table 4
### Post-treatment Effects of Working Memory Training

<table>
<thead>
<tr>
<th>Measure</th>
<th>F(1, 24)</th>
<th>Group p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criterion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS Digit Span: Forwards</td>
<td>1.29</td>
<td>.30</td>
<td>.11</td>
</tr>
<tr>
<td>WAIS Digit Span: Backwards</td>
<td>7.45</td>
<td>.004</td>
<td>.43</td>
</tr>
<tr>
<td>WAIS Digit Span: Sequencing</td>
<td>.35</td>
<td>.71</td>
<td>.034</td>
</tr>
<tr>
<td>CANTAB Spatial Span: Forwards</td>
<td>1.11</td>
<td>.35</td>
<td>.10</td>
</tr>
<tr>
<td>CANTAB Spatial Span: Backwards</td>
<td>4.60</td>
<td>.023</td>
<td>.32</td>
</tr>
<tr>
<td>WRAML Finger Windows: Forwards</td>
<td>2.04</td>
<td>.16</td>
<td>.17</td>
</tr>
<tr>
<td>WRAML Finger Windows: Backwards</td>
<td>2.06</td>
<td>.15</td>
<td>.17</td>
</tr>
<tr>
<td><strong>Near-Transfer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CANTAB Spatial Working Memory: Strategy Score</td>
<td>1.36</td>
<td>.28</td>
<td>.12</td>
</tr>
<tr>
<td>CANTAB Spatial Working Memory: Between Errors</td>
<td>1.99</td>
<td>.16</td>
<td>.17</td>
</tr>
<tr>
<td>CANTAB Pattern Recognition Memory</td>
<td>.27</td>
<td>.76</td>
<td>.03</td>
</tr>
<tr>
<td>Kahneman’s Addition Task</td>
<td>.414</td>
<td>.67</td>
<td>.04</td>
</tr>
<tr>
<td><strong>Far-Transfer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASRS</td>
<td>1.64</td>
<td>.22</td>
<td>.15</td>
</tr>
<tr>
<td>CFQ</td>
<td>3.46</td>
<td>.06</td>
<td>.348</td>
</tr>
<tr>
<td>BDEFS</td>
<td>.04</td>
<td>.961</td>
<td>.004</td>
</tr>
</tbody>
</table>
Table 5

Confidence Intervals for Post-treatment Effects of Working Memory Training

<table>
<thead>
<tr>
<th></th>
<th>Standard-length compared with shortened-length training</th>
<th>Standard-length training compared with Waitlist Control</th>
<th>Shortened-length training compared to Waitlist Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean difference</td>
<td>CI</td>
<td>Mean difference</td>
</tr>
<tr>
<td><strong>Criterion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS Digit Span: Forwards</td>
<td>1.25</td>
<td>-1.54, 4.04</td>
<td>-1.95</td>
</tr>
<tr>
<td>WAIS Digit Span: Backwards</td>
<td>1.75</td>
<td>-.50, 4.00</td>
<td>3.85*</td>
</tr>
<tr>
<td>WAIS Digit Span: Sequencing</td>
<td>.87</td>
<td>-.146, 3.19</td>
<td>.69</td>
</tr>
<tr>
<td>CANTAB Spatial Span: Forwards</td>
<td>.92</td>
<td>-.51, 2.35</td>
<td>.79</td>
</tr>
<tr>
<td>CANTAB Spatial Span: Backwards</td>
<td>2.11*</td>
<td>.59, 3.63</td>
<td>1.56*</td>
</tr>
<tr>
<td>WRAML Finger Windows: Forwards</td>
<td>2.27</td>
<td>-.124, 5.78</td>
<td>2.94</td>
</tr>
<tr>
<td>WRAML Finger Windows: Backwards</td>
<td>4.79</td>
<td>-.61, 10.20</td>
<td>3.85</td>
</tr>
<tr>
<td><strong>Near-Transfer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CANTAB Spatial Working Memory: Strategy Score</td>
<td>-2.59</td>
<td>-7.09, 1.91</td>
<td>-3.10</td>
</tr>
<tr>
<td></td>
<td>-9.51</td>
<td>-21.48, 2.45</td>
<td>-9.57</td>
</tr>
<tr>
<td><strong>Between Errors Score</strong></td>
<td>-.13</td>
<td>-24.26, 24.00</td>
<td>7.11</td>
</tr>
<tr>
<td>CANTAB Pattern Recognition Memory</td>
<td>.37</td>
<td>-1.72, 2.45</td>
<td>-.51</td>
</tr>
<tr>
<td>Kahneman’s Addition Task</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASRS</td>
<td>3.33</td>
<td>-2.64, 9.30</td>
<td>-1.49</td>
</tr>
<tr>
<td>CFQ</td>
<td>-.82</td>
<td>-8.79, 7.15</td>
<td>-9.29*</td>
</tr>
<tr>
<td>BDEFS</td>
<td>-.84</td>
<td>-7.09, 5.42</td>
<td>-.39</td>
</tr>
</tbody>
</table>

Figure 1
Flow of participants through the trial

Eligibility Assessed
(n=57)

Excluded
(n=19)
- Age (n=5)
- Comorbid psychiatric diagnosis (n=11)
- No longer a student (n=3)

Randomized
(n=38)

Standard-Length Training Group
(n=18)
- Received allocated intervention and provided post-test data (n=8)
- Left study prior to receiving intervention and did not provide post-test data (n=4)
  - Family emergency or illness (n=3)
  - Did not provide a reason (n=3)

Shortened-Length Training Group
(n=8)
- Received allocated intervention and provided post-test data (n=6)
- Left study prior to receiving intervention and did not provide post-test data (n=1)
  - Did not provide a reason (1)

Waitlist Control Group
(n=12)
- Received allocated intervention and provided post-test data (n=9)
- Received allocated intervention but did not provide post-test data (n=1)
- Left study prior to receiving intervention and did not provide post-test data (n=1)
  - Did not provide a reason (1)
- Left study mid-way through intervention and did not provide post-test data (n=1)
  - Did not provide a reason (1)

Figure 2. Post-test Effects of Working Memory Training on Manipulation of Auditory-Verbal Information
Figure 3. Post-test Effects of Working Memory Training on Manipulation of Visuo-Spatial Information
Figure 4. Post-test Effects of Working Memory Training on Everyday Cognitive Failures
References


Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65-94.


Appendix

Description of WM tasks in the Cogmed WM Training Program

1. Reproducing a light sequence in a visuo-spatial grid

Lamps arranged in a four-by-four grid are displayed. Participants watch several lights go on and then reproduce the same sequence.

2. Reproducing a light sequence in a rotated grid

A rotating version of the visuo-spatial grid task described above. After the sequence of lights goes on, the grid panel rotates 90 degrees clockwise and participants reproduce the sequence in the panel’s new position.

3. Indicating numbers in reverse order

A keyboard with numbers is displayed and then digits are read aloud. Participants respond by indicating the same numbers but in reverse order.

4. Indicating numbers in reverse order with a non-visible keyboard

This is a variation of the previous task; the keyboard with numbers is not displayed as the digits are read aloud, becoming visible only when participants respond.

5. Identifying letter positions in a sequence

Letters are read aloud, one at a time. Participants have to remember the letters in the order in which they are read. A row of lights is then visible and a flashing light cues the participant to indicate the letter that was read in the sequence. For example, if light number 3 lights up, then participants report the 3rd letter that they have just heard.

6. Identifying letter sequences

A sequence of letters is first read out loud. Then the participants are presented with three options letters and must select the letter that was presented. For example, if the letter
sequence “D, P, E” is presented. The participant must first select the letter D from one of three options before following a similar procedure with P and E.

7. Reproducing a light sequence in a rotating circle

A set of lamps are arranged in a rotating circle. Participants watch several lights go and then reproduce the same sequence, even though the lamps are constantly shifting positions.

8. Reproducing a light sequence

A number of moving circles randomly appear. Participants must reproduce the order in which the circles appeared.

9. Reproducing a sequence of moving shapes

A number of moving shapes light up in succession. The participants must remember the sequence in which the shapes lit up.

10. Reproducing a light sequence in a 3D visuo-spatial grid

Lights are symmetrically positioned in a 3D ‘room’ with 20 segments. Participants watch several lights go on and then reproduce the same sequence.

11. Reproducing a light sequence in a 3D visuo-spatial cube

Lights are symmetrically positioned in a 3D Cube with 12 segments. Participants watch several lights go on and then reproduce the same sequence.

12. Reproducing a sequence of numbers on a visual grid

A four by four grid is presented with 16 latches. A sequence of latches is opened displaying a set of numbers. The participant must sort the numbers by clicking on the latch that contained the numbers in numerical order.