ORIGINAL ARTICLE

Will Working Memory Training Generalize to Improve Off-Task Behavior in Children with Attention-Deficit/Hyperactivity Disorder?

Chloe T. Green · Debra L. Long · David Green · Ana-Maria Iosif · J. Faye Dixon · Meghan R. Miller · Catherine Fassbender · Julie B. Schweitzer

Abstract Computerized working memory and executive function training programs designed to target specific impairments in executive functioning are becoming increasingly available, yet how well these programs generalize to improve functional deficits in disorders, such as attention-deficit/hyperactivity disorder (ADHD), beyond the training context is not well-established. The aim of this study was to examine the extent to which working memory (WM) training in children with ADHD would diminish a core dysfunctional behavior associated with the disorder, "off-task" behavior during academic task performance. The effect of computerized WM training (adaptive) was compared to a placebo condition (nonadaptive) in a randomized, double-blind, placebo-controlled design in 26 children (18 males; age, 7 to 14 years old) diagnosed with ADHD. Participants completed the training in approximately 25 sessions. The Restricted Academic Situations Task (RAST) observational system was used to assess aspects of off-task behavior during the completion of an academic task. Traditional measures of ADHD symptoms (Conners' Parent Rating Scale) and WM ability (standardized WM tests) were also collected. WM training led to significant reductions in off-task ADHD-associated behavior on the RAST system and improvement on WM tests. There were no significant differences between groups in improvement on parent rating scales. Findings lend insight into the generalizability of the effects of WM training and the relation between deficits in WM and off-task behavioral components of ADHD. These preliminary data suggest WM training may provide a mechanism for indirectly altering academic performance in children with ADHD.

Keywords Attention · academic behavior · treatment · RAST · children · ADHD · cognitive training.

C. T. Green · J. F. Dixon · M. R. Miller · C. Fassbender ·
J. B. Schweitzer (✉)
Department of Psychiatry and Behavioral Sciences and MIND Institute, University of California Davis School of Medicine, Sacramento, CA 95817, USA
e-mail: julie.schweitzer@ucdmc.ucdavis.edu

C. T. Green
e-mail: ctgreen@berkeley.edu

C. T. Green · D. L. Long
Department of Psychology, University of California, Davis, Davis, CA, USA

D. Green
43 Quail Court Suite 111,
Walnut Creek, CA 94596, USA

A.-M. Iosif
Department of Public Health Sciences, University of California, Davis,
Davis, CA, USA

M. R. Miller
Department of Psychology, University of California, Berkeley, Berkeley, CA, USA

Present Address:
C. T. Green
School Psychology Doctoral Program,
University of California, Berkeley,
CA, USA

Introduction

Nonpharmacological interventions continue to be of great interest to parents of children with attention-deficit/hyperactivity disorder (ADHD). Training caregivers in behavioral procedures for ADHD is well-supported [1–3]; however, often the treatment does not generalize beyond the clinical

setting. Newer behavioral treatments are attempting to address this issue [4], but providing evidence of generalization remains a challenge for most nonpharmacological therapies for ADHD. The effects of a relatively new treatment for ADHD, working memory training (WMT), merits further examination as to its ability to generalize beyond the direct treatment context, because initial research in typically developing adults suggests that WMT directly affects functioning of brain regions, such as the prefrontal cortex (PFC) and posterior parietal lobe [5–10], as well as dopamine (a neurotransmitter altered in ADHD) receptor binding [11].

Interventions that target working memory (WM) deficits are particularly relevant to ADHD due to the presumed central role WM plays in learning and reasoning. Critical to classroom learning, WM is the function of actively holding in mind and manipulating information relevant to a goal [12]. WM is a core cognitive function essential for academic performance and achievement, goal attainment, and following rules [13–15]. WM was viewed as a limited capacity system, and thus children with impaired WM often experience overload during learning activities, hindering their ability to sustain attention and stay on task [13]. Neuropsychological testing reveals moderate-to-marked impairment in WM in children with ADHD, both in initiating WM and in the limits of storage, particularly in the visuospatial domain [16–24]. These WM deficits may contribute significantly to inefficient learning, behavioral problems, executive dysfunction, and the eventual underachievement that children with ADHD often experience [25–29].

WM was traditionally believed to be a fixed genetic trait, the capacity of which could not improve through practice [30]; however, recent studies show that WM can be improved through intensive training [31–36]. Double-blind, placebo-controlled studies suggest that participation in 25 sessions of intensive computerized training significantly improves WM performance in children with ADHD [31–33]. Enhanced performance on nontrained WM tasks was seen in both visuospatial and verbal WM, and persisted at 3 months and at 6 months after the end of the training period, suggesting that improvements after treatment are durable after an extent of time [31, 33]. Holmes et al. [37] recently compared the effect of WMT to the effects of stimulant medication on short-term memory and WM in 25 children with ADHD. The dose of medication was prescribed by the treatment provider of the children. The authors found that the training improved verbal and visuospatial WM and short-term memory functioning, whereas stimulant medication only resulted in improvement in visuospatial WM. Furthermore, a follow-up analysis demonstrated that gains in performance for visuospatial short-term memory, verbal WM, and visuospatial WM persisted 6 months after the training concluded.

A recent study of 52 children with ADHD in a wait list, controlled trial found significant improvement on parent-rated measures of executive functioning, the Behavior Rating Inventory of Executive Function (BRIEF), immediately after training and at a 4-month follow-up [38]. However, there were no significant changes in teacher ratings in this study, nor were there any changes in a previous study that included teacher ratings [33]. These studies demonstrated partial evidence for the generalization of the training on more global functioning on parent (but not teacher) ADHD ratings scales, improved response inhibition functioning, and complex reasoning, as measured by the Raven's Progressive Matrices [32, 33]. Thus, there is some evidence that computerized cognitive WM training may affect broader executive functioning.

Neuroimaging studies reveal abnormal PFC and parietal cortex functioning in children with ADHD during WM tasks [5–9, 39, 40]. Studies of the neural mechanism of action of WMT in 2 modestly sized functional magnetic resonance imaging studies in healthy adults ($n=3$ and $n=8$) in Olesen et al. [35] and ($n=3$) in Westerberg and Klingberg [41] suggest that WMT increases brain activity in the PFC and parietal cortex [35], regions frequently associated with WM processing and executive functioning in general [42, 43]. A functional magnetic resonance study of training in ADHD using a broad package of exercises [44], focusing on response inhibition and selective attention rather than WM, found increased frontal, middle temporal, and cerebellar activity after training. Theoretically, improvement in dopamine functioning, which is altered in ADHD and is targeted by the most efficacious ADHD pharmacologic treatments [45], would also enhance the generalization of training effects. A recent positron emission tomography report of 13 healthy adults [11] found increased density of cortical dopamine D1 receptor binding in the PFC and parietal cortex associated with WM training. Thus, training may improve behavioral and cognitive impairments of ADHD related to disturbances of the dopamine system [11, 45]. These neuroimaging findings suggest that WM training may have the potential for improved generalization by targeting core weaknesses involved in executive functioning in ADHD populations.

The goal of the current study was to explore whether the effects of computerized WMT would generalize to increase "on-task" behavior during performance of an academic task. "Off-task" behavior during homework and "in-seat" academic assignments is one of the most common reasons parents of children with ADHD seek evaluations and treatment. Research shows that off-task behavior correlates with impaired WM in children with and without ADHD [22]. We used the Restricted Academic Setting Task (RAST), a simulated classroom setting task and observational system that is a good indicator of the behavioral response of children to the pharmacological treatment effects for ADHD [46–49]. The RAST quantifies the frequency and type of the off-task

behavior of a child during academic task performance that may not be salient in teacher observations. Elevated frequencies of off-task behavior on the RAST correlate with elevated scores for externalizing behaviors on teacher ratings of ADHD [50] and actometer ratings [51]. The use of a laboratory measure permitted us to obtain objective outcomes that are not susceptible to the measurement and informant bias that would be encountered in a classroom situation. We selected a task related to performance during an academic task because ADHD is associated with negative outcomes in academic achievement, including lower reading and math achievement scores [52–56], lower high school completion rates [57], and an alarmingly high use of special education services [14]. In the current project, we hypothesized that 25 sessions of intensive, adaptive WM training would lead to improved on-task behavior during a simulated academic task (i.e., RAST), traditional WM measures, and parent ratings of ADHD.

Method

Participants

The research protocol, consent, and assent were approved by the University of California Davis Institutional Review Board. Before entering the study, written consent was obtained from the parent or legal guardian of the participant, and verbal assent for children under 13 or written assent for children 13 years and older. The study was conducted in accordance with the Declaration of Helsinki [58]. Participants were recruited through the subject tracking and recruitment system at the institute, by psychologists and psychiatrists, and through flyers or advertisements that were posted locally and to a support group.

Thirty-two participants were pre-screened for the study, thirty were randomized, with 1 participant withdrawing from the placebo and 3 from the adaptive training condition. Thus, there were a total of 26 participants completing the study, including 12 in the active training condition and 14 in the placebo control condition. For the final sample, participants ranged in age from 7 to 14 years of age (active training mean age, 9.9 years; placebo mean age, 9.6 years) with no significant differences between groups in either age or gender composition (see Table 1 for full demographic and clinical summaries).

Each child and parent or guardian participated in an evaluation session to determine eligibility for the study and to establish a baseline score on the preintervention and postintervention measures. Final determination of eligibility for the study was conducted by a licensed, Ph.D., psychologist, based on the following criteria and administered in the following order: 1) positive phone screening for ADHD and no exclusions; 2) positive for ADHD using the Diagnostic

Table 1 Sample Demographic and Clinical Characteristics

	Placebo (n=14)	Treatment (n=12)	Total (n=26)
Demographic characteristics			
Gender			
Boys	9 (64 %)	8 (67 %)	17 (65 %)
Girls	5 (36 %)	4 (33 %)	9 (35 %)
Age	9.6±2.6	9.9±1.8	9.7±2.2
Race			
Caucasian	9 (64%)	8 (67%)	17 (65%)
African American	1 (7%)	1 (8%)	2 (8%)
Asian	3 (21%)	3 (25%)	6 (23%)
American Indian	1 (7%)	0 (0%)	1 (4%)
Ethnicity			
Non-Hispanic or Latino	12 (86%)	9 (75%)	21 (81%)
Hispanic, of Spanish origin or Latino	1 (7%)	2 (17%)	3 (12%)
Unknown	1 (7%)	1 (8%)	2 (8%)
Clinical Characteristics			
WASI FSIQ	105.4±14.2	107.0±12.3	106.2±13.1
ADHD diagnosis			
Combined	6 (43 %)	5 (42 %)	11 (42 %)
Inattentive	8 (57 %)	5 (42 %)	13 (50 %)
Hyperactive/Impulsive	0 (0 %)	2 (17 %)	2 (8 %)
Using medication	2 (14 %)[*]	8 (67 %)[*]	10 (38 %)

Data are summarized as mean±SD for the continuous variables and frequency (%) for the categorical ones.

*Fisher's exact test (p=0.01)

ADHD=attention-deficit/hyperactivity disorder; FSIQ=Full Scale Wechsler Abbreviated Scale of Intelligence; WASI=Wechsler Abbreviated Scale of Intelligence

and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision criteria for ADHD (any subtype) using the Diagnostic Interview for Children and Adolescents — Parent version; 3) T-score of at least 65 on any of the ADHD Diagnostic and Statistical Manual scales of the Conners' Parent Rating Scale-Revised (CPRS-R); and 4) no history of mental retardation (i.e., IQ > 70) on the Wechsler Abbreviated Scale of Intelligence (WASI). Further information was gathered on each child to document relevant characteristics, including medication history and presence of a learning disability. Learning disability status was assessed using standardized achievement tests (Woodcock Johnson, Third Edition) to assess academic functioning and the presence of an academic learning disability. A learning disability was diagnosed as the presence of a significant IQ achievement discrepancy (defined as 1.5 SD) and also low achievement (i.e., standard score <80 on a Woodcock Johnson, Third Edition subtest) test. Parents completed a demographic form providing general information regarding the participant (e.g., age, education; parents' age and education). Conners' Teacher Rating Scale-Revised (CTRS-R) were obtained to supplement parent ratings for establishing the diagnosis when available; however, because many teachers were not available to complete the ratings (e.g., many children were recruited during the summer break) or did not return the scales; there were an insufficient number of scales to be used as an outcome measure). Children with a diagnosis of the following were excluded: severe mental illness (i.e., psychotic, bipolar, or major depressive disorder, by history or clinical interview) or autism spectrum disorders. Inclusion in the study also required reliable access to the internet, and parents and children who were fluent in English. Parents were informed at the time of consent that we expected them not to change their child's medication (or start medication if they were not taking medication) or engage in any other treatment during the course of the study. This was deemed feasible because the intervention phase was relatively short. Parents were instructed that treatment changes could occur if it became medically necessary in emergency conditions.

Study Design

After screening, participants were randomly assigned to either the treatment (adaptive WM training) or to the control placebo (nonadaptive) condition using a computerized generated random number sequence. Throughout the study, both the participants and researchers remained blind to the condition assigned to the participant. The coach, who had contact with participants during the study, was not involved in the random assignment, in testing on any outcome or screening measures, or in the data analytic components of the study.

Outcome Measures

The primary measure of efficacy was change between the RAST score at the pre-WMT baseline assessment and the RAST score during the post-WMT assessment. The RAST score is used to assess the effects of pharmacological treatment, as it permits an objective measure of behavior in a controlled laboratory setting [59]. The RAST provided information regarding the frequency of off-task behaviors during performance on an academic task on 5 areas (i.e., off-task, "out-of-seat," "fidgets," "vocalizes," and "plays with object"). The RAST is sensitive to moment-to-moment off-task behavior that a teacher or parent may be unlikely to detect.

The RAST was administered before and after training with participants videotaped through a 1-way mirror during the session. At the beginning of the task, the child was asked to sit at a table and play independently with a game or toy of his or her choice. After 5 minutes, the researcher entered the room, moved the game or toy to the side and told the child to complete a series of academic worksheets (e.g., math problems) for 15 minutes. The academic worksheets were at least 1 grade level below the current grade level of the child or his or her ability (based on achievement testing). Thus, the aim of the worksheets was to absorb the child's attention without being overly demanding from an academic perspective. Before leaving the room, the researcher instructed the child not to leave his or her choice seat or to touch any of the games or toys. The observer recorded (tallied) the occurrences of the following behaviors every 30 seconds: off-task (looks away from paper), out-of-seat (leaves chair), fidgets (repetitive purposeless motion), vocalizes, and plays with object ("touches any object in the room unrelated to the task") [60].

A single rater scored the observations of all the participants; the rater was trained by an experienced RAST coder on how to code the behaviors on the RAST, which the rater did for each participant. Then 20 % of the RAST data were scored by a second trained RAST coder to verify reliability in the RAST scoring. Both coders were blind to group membership. The agreement between the 2 raters was excellent across all categories with kappa at least 0.95 in all categories: off-task=0.95, plays with object=1.00, out-of-seat=0.97, fidgets=0.96, and vocalizes=0.96.

The working memory Index (WMI) from the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV) assessed change in WM performance. This index includes the Digit-Span and Letter-Number Sequence subtests. The CPRS-R served as an additional outcome measure.

Intervention

Each participant was required to complete 90 trials of WM Cogmed (http://www.cogmed.com) tasks a day for 25 days.

Typically, this could be accomplished within 40 minutes per day but some participants may have finished the session in as short as 25 minutes. Each participant received the program either on a CD or via e-mail containing a software download link to the program. The children used the program on their personal computers at home. Researchers instructed parents to supervise his or her child as they each performed the tasks.

The Cogmed training includes 10 verbal and visuospatial WM span tasks. Some tasks are both auditory and visual in nature, requiring cross-modal processing. During each session, participants choose 6 of 8 presented tasks. During the active training condition, for each task, the difficulty level automatically adjusts to match the WM span of the child. The difficulty level gradually increased when the child answered correctly on consecutive trials and decreased when the child answered incorrectly on consecutive trials. The tasks require good attention and tracking ability, as some stimuli move on the screen. In the placebo control-nonadaptive condition, the same tasks were used, but the difficulty level remained low (2 or 3 numbers/items at maximum) throughout all of the training sessions. These similar versions controlled for nonspecific effects of the training procedure and enabled us to attribute any improvement in ADHD-related behaviors and WM exclusively, due to the effect of training WM.

Coaching was used to enhance compliance with completing the sessions for both the placebo and training condition. The same licensed clinical psychologist coached all participants during each week of training on the telephone at least once a week. Coaching involved answering questions regarding the use of the computer program and troubleshooting software issues, general feedback for the use of the program, and addressing parental concerns of how to engage their child in the training protocol. Coaching was kept to a minimum so as to reduce any possible differences between groups in amount or type of feedback. The coach had access via the internet to the frequency of use and performance on the computer tasks. The software is programmed to provide participants with auditory and visual feedback after each trial, which informs them of whether they were correct or incorrect. The program also included an optional reward game at the end of each training session that most subjects reported enjoying. The feedback and the rewarding game at the end of the session were provided for participants in both groups.

Statistical Analysis

Group differences in demographic and clinical characteristics were assessed using Fisher's exact tests for the categorical variables and two sample t tests (or appropriate nonparametric tests if the data violated the assumption of normality) for the continuous ones. The RAST provides continuous scores on each subscale; however, the distributions for vocalizes, plays with object, and out-of-seat were strongly skewed to the right and had a spike at 0, thus neither the normal distribution, nor count distributions were applicable to these data. Therefore, we dichotomized each of them into behavior being present or absent for the entire 15-minute duration of the task. A Generalized Linear Models [61] approach was used to analyze all dependent variables because it is appropriate for both outcomes that were not normally distributed (e.g., vocalizes, plays with object, and out-of-seat) and for those outcomes normally distributed (e.g., off-task, fidgets, and WMI). This method takes into account the correlated structure of the data due to repeated assessments in time and allows for missing observations. The goals of the statistical models were to estimate change in the outcome variables for the course of the study and to test whether the treatment group was related to either initial level or rate of change in the dependent variables. Thus, for each of the dependent variables, we first fitted a model with a main effect for the treatment group (placebo or active WM training), time (pre- and post-intervention), and their interaction, a term for medication use, and a random effect for the child, to account for the correlation due to repeated measures on the same individual. A significant group by time interaction indicates that the change in outcome from pre- to post-intervention differs between the 2 treatment groups. After fitting this initial model, we examined another set of models, in which a term for mathematical ability was added and tested to investigate its effect on the outcome variables. A variable was considered to be a significant predictor in the model if its significance level exceeded 0.05. All analyses were implemented using PROC GENMOD and GLIMMIX in SAS (SAS Institute, 2002–2010, Carey, NC, USA), version 9.2.

Results

The treatment group included a significantly greater number of participants with actively prescribed medication for ADHD ($n=8$) than the comparison group ($n=2$) using Fisher's exact test ($p=0.01$). Preliminary analyses compared the baseline levels of RAST behaviors and WMI between the medicated and nonmedicated children. Use of medication was not associated with better performance on any of these measures. Furthermore, medication use was included as a covariate in all statistical analyses, but it did not have a significant effect on any of the RAST behaviors or WMI. The groups had similar full scale IQ, as assessed by the Wechsler Abbreviated Scale of Intelligence, with a mean of 107.0±12.3 SD) for the active treatment group and a mean of 105.4±14.2 SD) for the placebo group. There were an insufficient number of subjects from each ADHD subtype to explore, if there were any differences due to subtype in

response to training. There was 1 participant who was diagnosed with an academic learning disability ($n=1$, writing) who was in the placebo condition. Participants in the placebo group completed the training in a mean time of 23.57 days ±2.34 SD), and the adaptive training condition in a mean time of 24.33 days ±1.92 SD), with a range of training between 20 and 27 days.

Restricted Academic Setting Task

Table 2 summarizes RAST behaviors for the 2 groups: 1) pre-intervention and 2) post-intervention. As depicted in Fig. 1, the off-task behavior was similar in the 2 groups ($p=$ 0.29) pre-intervention, but the active treatment group had a sharp decline in looks away behavior post-intervention, whereas the placebo group maintained the pre-intervention levels. The interaction term between group and time was significant, and the difference in improvement between the 2 groups was 12.3 points (±4.6 SE; $p=0.01$). There were no significant training effects on fidgeting, with both groups maintaining the observed pre-intervention levels (difference in improvement 0.9; ±3.9 SE; $p=0.81$). Next, we analyzed the RAST subcategories occurring less frequently for both groups (the group plays with object, and the group vocalizes and out-of-seat) by treating them as binary outcomes. We found that training led to a reliable decrease in the plays with object behavior. The interaction term between group and time was significant (estimate, 3.4; $p=0.04$). Although children in the placebo group displayed similar rates of this behavior pre- and post-intervention, those in the active treatment had a significant drop post-treatment (with only 1 of 12 children showing this behavior after intervention). There were no significant
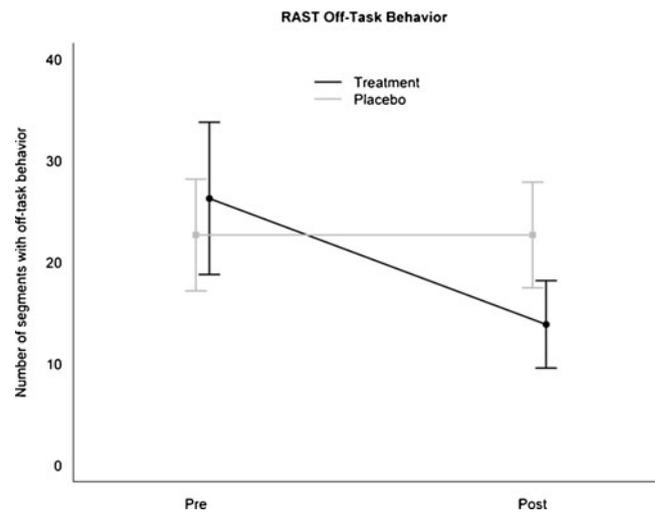
Fig 1 Average and 95 % confidence intervals for the number of 30-second intervals of children engaged in off-task behavior during the Restricted Academic Situations Task (RAST) pre- and post-intervention for the 2 treatment conditions

differences in either baseline levels or changes as a result of training for both out-of-seat and vocalizes behaviors.

Nontrained WMT

To examine whether training effects transferred to nontrained WM tasks, we analyzed the effect of the training on the WISC WMI. There were no significant differences between the 2 groups pre-intervention. As hypothesized, there was a significant interaction between group and time (see Table 3). As shown in Fig. 2, this interaction was driven by a significant improvement on WM tests in the treatment group that was not

Table 2 Summary for the Behaviors in the Restricted Academic Setting Task, Pre- and Post-Intervention

	Treatment group				Difference in improvement between active and placebo* Estimate (SE)	F Value[†]	p Value
	Placebo		Active Treatment				
	Pre-intervention	Post-intervention	Pre-intervention	Post-intervention			
Continuous outcomes, Mean±(SD)							
Off-task	22.7±9.5 SD	22.7±9.0 SD	26.2±11.8 SD	13.9±6.7 SD	12.3±4.6 SE	7.10	0.01
Fidgets	19.1±9.2 SD	18.5±8.8 SD	18.1±11.8 SD	16.5±8.8 SD	0.9±3.9 SE	0.06	0.81
Binary outcomes,[‡] frequency (%)							
Out-of-seat	8 (57 %)	7 (50 %)	9 (75 %)	5 (42 %)	1.2±1.2 SE	0.97	0.33
Vocalizes	10 (71 %)	6 (43 %)	7 (58 %)	5 (42 %)	−0.6±1.2 SE	0.23	0.63
Plays with objects	8 (57 %)	7 (50 %)	9 (75 %)	1 (8 %)	3.4±1.5 SE	4.99	0.04

*All models controlled for use of medication

[†] All F statistics have 1, 24 degrees of freedom

[‡] The behavior in these categories was dichotomized into behavior present or absent, as these variables had an excessive number of zeros. Entries in table indicate number (%) of children displaying the behavior

Table 3 Summary for the WISC WM Index, Pre- and Post-intervention

	Treatment group mean (SD)				Difference in improvement between active and placebo* Estimate (SE)	F Value[†]	p Value
	Placebo		Active treatment				
	Pre-	Post-	Pre-	Post-			
WISC WMI composite[‡]	90.0±10.1 SD	88.0±13.5 SD	95.9±10.6 SD	102.3±9.1 SD	8.4±3.3 SE	6.67	0.02

WISC=Wechsler Intelligence Scale for Children-IV; WMI=working memory index

*All models controlled for use of medication

[†] F statistic has 1; 24 degrees of freedom

[‡] One observation missing in the placebo group

present in the placebo group (difference in improvement between active and placebo=8.4±3.3 SE; p=0.02).

Rating Scales

The placebo group was rated an average t-score of 78.1±8.3 SD at baseline and 70.9±8.3 SD post-treatment on the ADHD Index of the CPRS-R. The average t-score on the ADHD Index of the CPRS-R in the WM training group was 72.5± 7.5 SD at baseline and 67.0±12.2 SD after treatment. Thus, the parent ratings on the CPRS-R ADHD Index decreased significantly with time for both the placebo and the active training condition, but the improvement was similar in the 2 groups (estimated difference=− 1.3; ±4.0 SE; p=0.74).

Discussion

This placebo-controlled, randomized, and double-blind study of WM training for children with ADHD is the first,

Fig 2 Average and 95 % confidence intervals for the Wechsler Intelligence Scale for Children (WISC) Working Memory Index (WMI) scaled score pre- and post-intervention for the 2 treatment conditions

to our knowledge, that demonstrates improved performance in an ecologically valid laboratory measure of observable ADHD-associated behaviors. Our primary measure (i.e., the RAST) quantifies subtle improvements in performance. Training had the greatest effect on the category in which the highest rate of behavior occurred, specifically, the off-task category, which is measured by whether the child looks away from the worksheets. This measure might be most related to attention. The influence of training on these categories, particularly the off-task category, is significant because inattention is often considered to be a primary issue in both the inattentive and combined subtypes of ADHD, and is related to academic functioning. We also saw a reduction in the plays with objects category. There may have been some overlap between the off-task and plays with objects categories, although it is conceivable that a child could hold or touch a nontask related object while maintaining attention to the worksheets.

Although prior studies have found that WM training led to significant improvements on the CPRS-R [33, 38] short form, we found that both our placebo and active treatment group improved, without significant differences between the groups, when we administered the CPRS long form. One primary difference between our study and the Klingberg et al. [33] study is that the majority of our participants in the adaptive training condition were prescribed stimulant medication, whereas none of the participants in the Klingberg et al. [33] study were taking stimulant medication. Thus, it is possible that there was less room to show improvement on the rating scale for our participants. Consistent with this, our baseline CPRS-R ratings for children in the treatment group were lower than those in the placebo group. However, we do note that the Beck et al. [38] study was similar to ours in that many of their participants were prescribed stimulant medication. Other differences between our sample and those of Beck et al. [38] and Klingberg et al. [33] may have also affected the sensitivity to detect change on the parent rating scales. For example, those samples included participants with a higher preponderance of comorbid disorders and the inattentive subtype, whereas our sample did not. The
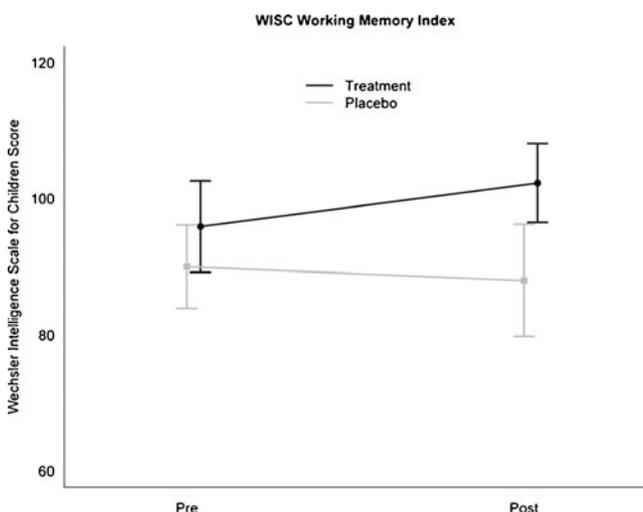
Klingberg et al. [33] study found a significant improvement on the Oppositionality scale of the Conners' Scale, yet because we did not have a high rate of oppositional behavior in our sample, we would be less likely to show improvement on that measure in our participants.

Future studies could measure the effect of the training on off-task behavior of the child in his or her own classroom (see the Achenbach System of Empirically Based Assessment Direct Observation Form); however, there would be less standardization between the data acquisition periods in comparison to the RAST procedure. Limitations in teacher ratings also may exist, as the attention of teachers, dispersed among a group of students, is often selectively pulled toward behavioral management. Thus, teachers may not be sensitive in detecting positive changes in on-task behavior and may be less objective in light of having already formed a general impression of the behavioral patterns of the child. The absence of measurable effects of WM training on teacher ratings in previous studies is a notable limitation of the training.

The training did lead to improvement on the WISC-IV WMI, a widely used standardized WM measure. This improvement was only observed in the training group. Subtests that comprise the WMI, the Letter-Number Sequencing and Digit Span, are similar to the training exercises, and thus, improvement on this measure was not unexpected. We recommend that future studies assess academic performance and achievement, planning, goal attainment, and following rules [13–15] as outcome measures, as they are all associated with WM performance and represent challenges for persons with ADHD.

The training program (i.e., Cogmed) that we used is associated with increases in PFC and parietal lobe activity, and in dopamine D1 receptor binding in healthy adults [11]. Both PFC and dopamine are associated with the general category of executive functioning and learning, and may be responsible for the generalization to nontrained behavior. The training also includes delivery of rewards for correct performance, which likely acts on the basal ganglia, although both training and placebo groups would have received rewards. Stimulant treatment, such as methylphenidate, is known to act directly on brain regions associated with disrupted dopamine reward pathways in ADHD [62, 63] and may be responsible for the global effects seen with stimulant treatment in ADHD. Understanding and comparing the effects of treatments that directly manipulate dopaminergic functioning via pharmacological approaches *versus* delivering of conditioned rewards as feedback within the training system have not been conducted. Such comparisons have the potential to improve how nonpharmacological treatments are used for disorders associated with reward dysfunction, such as ADHD [64] or schizophrenia [65].

Although our study found a significant effect of WM training on ecologically valid measures, several limitations should be noted. A greater number of participants had been prescribed stimulant medication in the training group than in the placebo groups, and this may be a potentially confounding variable to consider. Thus, we are not able to determine if an interaction between medication and WM training may have influenced the findings. Furthermore, although parents of participants were told not to change their treatment regimen, 1 participant in the treatment condition was on medication at pre-testing and discontinued medication during the treatment and at post-testing. Participants in the placebo condition likely received more positive feedback than did participants in the training condition. Future studies could use a yoked design with participants receiving an equivalent degree of feedback. Our sample size was also quite modest and replications in larger samples are warranted. Comparisons between dropout rates of participants for training exercises *versus* another treatment or placebo should be explored, as we noted a slightly higher dropout rate for the training ($n=3$) *versus* placebo ($n=1$) groups.

Interventions that have the potential to improve WM and on-task behavior in academic-related performance are critical to ADHD because of their pivotal importance to success in academic settings. Given that available services for child psychiatric treatment are limited in many areas, and, when they are available, they are frequently underused by families who require treatment of a child with ADHD, the potential availability of computerized WM training in a school setting is especially appealing. Nonadherence with treatment recommendations in ADHD ranges from 20 to 90 % [66, 67], with nonadherence by parents partly due to concerns regarding the possible negative effects of medication. Treatments that may reduce the dose for medication can fill an important niche for parents who are averse to pharmacological interventions. How well WMT performs in comparison to traditional treatments over extended periods for ADHD at this time, however, is not known.

Subsequent studies should address the potential for training to enhance WM in larger samples of ADHD. Future research should also isolate the effect of medication status on the treatment effect to investigate if there is a synergistic effect of medication and WM training. Direct comparative efficacy studies comparing medication and other treatments for ADHD (e.g., parent training) to WMT should be conducted to further guide research and practitioners. Furthermore, the underlying neural changes associated with WM training are not well understood. Additional neuroimaging studies are needed to examine the effects of training on brain activation patterns, the neurochemical effects of training, and if ADHD subtype or comorbidity affects the efficacy of WMT programs.

Finally, future studies should examine the effects of combined WM training plus ADHD behavioral coaching

aimed at skill building and maintenance of skills to address specific impairments in ADHD. The current study suggests that brief, WMT can generalize to improve nontrained ADHD-related impairments; however, future training programs may be improved by developing modules that can be individually tailored to target additional impairments that children and adults with ADHD experience.

Required Author Forms Disclosure forms provided by the authors are available with the online version of this article.

References

1. Chronis AM, Jones HA, Raggi VL. Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. Clin Psychol Rev 2006;26:486–502.

2. Pelham WE Jr., Wheeler T, Chronis A. Empirically supported psychosocial treatments for attention deficit hyperactivity disorder. J Clin Child Psychol 1998;27:190–205.

3. Anastopoulos AD, Guevremont D, Shelton T, DuPaul GJ. Parenting stress among families of children with attention deficit hyperactivity disorder. J Abnorm Child Psychol 1992;20:503–520.

4. Abikoff H. ADHD psychosocial treatments: generalization reconsidered. J Atten Disord 2009;13:207–210.

5. Hale TS, Bookheimer S, McGough JJ, Phillips JM, McCracken JT. Atypical brain activation during simple & complex levels of processing in adult ADHD: an fMRI study. J Atten Disord 2007;11:125–140.

6. Krauel K, Duzel E, Hinrichs H, Santel S, Rellum T, Baving L. Impact of emotional salience on episodic memory in attention-deficit/hyperactivity disorder: a functional magnetic resonance imaging study. Biol Psychiatry 2007;61:1370–1379.

7. Schweitzer JB, Faber TL, Grafton ST, Tune LE, Hoffman JM, Kilts CD. Alterations in the functional anatomy of working memory in adult attention deficit hyperactivity disorder. Am J Psychiatry 2000;157:278–280.

8. Sheridan MA, Hinshaw S, D'Esposito M. Efficiency of the prefrontal cortex during working memory in attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2007;46:1357–1366.

9. Fassbender C, Schweitzer JB. Is there evidence for neural compensation in attention deficit hyperactivity disorder? A review of the functional neuroimaging literature. Clin Psychol Rev 2006;26:445–465.

10. Burgess GC, Depue BE, Ruzic L, Willcutt EG, Du YP, Banich MT. Attentional control activation relates to working memory in attention-deficit/hyperactivity disorder. Biol Psychiatry 2010;67:632–640.

11. McNab F, Varrone A, Farde L, et al. Changes in cortical dopamine D1 receptor binding associated with cognitive training. Science (New York) 2009;323:800–802.

12. Baddeley A. The fractionation of working memory. Proc Natl Acad Sci 1996;26:13468–13472.

13. Alloway TP, Gathercole SE, Elliott J. Examining the link between working memory behaviour and academic attainment in children with ADHD. Dev Med Child Neurol 2010;52:632–636.

14. Barkley RA. Attention Deficit hyperactivity disorder: a handbook for diagnosis and treatment, 3rd ed. New York: Guilford Press, 2006.

15. Gropper RJ, Tannock R. A pilot study of working memory and academic achievement in college students with ADHD. J Atten Disord 2009;12:574–581.

16. McLean A, Dowson J, Toone B, et al. Characteristic neurocognitive profile associated with adult attention-deficit/hyperactivity disorder. Psychol Med 2004;34:681–692.

17. Karatekin C, Asarnow RF. Working memory in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder. Psychiatry Res 1998;80:165–176.

18. Kempton S, Vance A, Maruff P, Luk E, Costin J, Pantelis C. Executive function and attention deficit hyperactivity disorder: stimulant medication and better executive function performance in children. Psychol Med 1999;29:527–538.

19. Kuntsi J, Oosterlaan J, Stevenson J. Psychological mechanisms in hyperactivity: I. Response inhibition deficit, working memory impairment, delay aversion, or something else? J Child Psychol Psychiatry 2001;42:199–210.

20. Mariani M, Barkley R. Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. Dev Neuropsychol 1997;13:111–129.

21. Martinussen R, Hayden J, Hogg-Johnson S, Tannock R. A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 2005;44:377–384.

22. Rapport MD, Bolden J, Kofler MJ, Sarver DE, Raiker JS, Alderson RM. Hyperactivity in boys with attention-deficit/ hyperactivity disorder (ADHD): a ubiquitous core symptom or manifestation of working memory deficits? J Abnorm Child Psychol 2009;37: 521–534.

23. Schweitzer JB, Hanford RB, Medoff DR. Working memory deficits in adults with ADHD: is there evidence for subtype differences? Behav Brain Funct 2006;2:43.

24. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol Psychiatry 2005;57:1336–1346.

25. Aronen ET, Vuontela V, Steenari MR, Salmi J, Carlson S. Working memory, psychiatric symptoms, and academic performance at school. Neurobiol Learn Mem 2005;83:33–42.

26. Gathercole SE, Pickering SJ. Working memory deficits in children with low achievements in the national curriculum at 7 years of age. Br J Educ Psychol 2000;70(part 2):177–194.

27. Rapport MD, Scanlan SW, Denney CB. Attention-deficit/hyperactivity disorder and scholastic achievement: a model of dual developmental pathways. J Child Psychol Psychiatry 1999;40:1169–1183.

28. Jarvis HL, Gathercole SE. Verbal and non-verbal working memory and achievements on national curriculum tests at 11 and 14 years of age. Educational and Child Psychology 2003;20:123–140.

29. Biederman J, Petty CR, Ball SW, et al. Are cognitive deficits in attention deficit/hyperactivity disorder related to the course of the disorder? A prospective controlled follow-up study of grown up boys with persistent and remitting course. Psychiatry Res 2009;170:177–182.

30. Kremen WS, Jacobsen KC, Xian H, et al. Genetics of verbal working memory processes: a twin study of middle-aged men. Neuropsychology 2007;21:569–580.

31. Holmes J, Gathercole SE, Dunning DL. Adaptive training leads to sustained enhancement of poor working memory in children. Dev Sci 2009;12:F9-F15.

32. Klingberg T, Forssberg H, Westerberg H. Training of working memory in children with ADHD. J Clin Exp Neuropsychol 2002;24:781–791.

33. Klingberg T, Fernell E, Olesen PJ, et al. Computerized training of working memory in children with ADHD–a randomized,

controlled trial. J Am Acad Child and Adolesc Psychiatry 2005;44:177–186.

34. Kurtz MM, Seltzer JC, Fujimoto M, Shagan DS, Wexler BE. Predictors of change in life skills in schizophrenia after cognitive remediation. Schizophr Res 2009;107:267–274.

35. Olesen PJ, Westerberg H, Klingberg T. Increased prefrontal and parietal activity after training of working memory. Nat Neurosci 2004;7:75–79.

36. Klingberg T. Training and plasticity of working memory. Trends Cogn Sci [Review]. 2010;14:317–324.

37. Holmes J, Gathercole SE, Place M, Dunning DL, Hilton KL, Elliott JG. Working memory deficits can be overcome: impacts of training and medication on working memory in children with ADHD. Appl Cogn Psychol 2010;24:827–836.

38. Beck SJ, Hanson CA, Puffenberger SS, Benninger KL, Benninger WB. A controlled trial of working memory training for children and adolescents with ADHD. J Clin Child Adolesc Psychol 2010;39:825–836.

39. Schweitzer JB, Lee DO, Hanford RB, et al. Effect of methylphenidate on executive functioning in adults with attention-deficit/hyperactivity disorder: normalization of behavior but not related brain activity. Biol Psychiatry 2004;56:597–606.

40. Fassbender C, Schweitzer JB, Cortes CR, et al. Working memory in attention deficit/hyperactivity disorder is characterized by a lack of specialization of brain function. PLoS One 2011;6:e27240.

41. Westerberg H, Klingberg T. Changes in cortical activity after training of working memory — a single-subject analysis. Physiol Behav 2007;92:186–192.

42. Olson IR, Berryhill M. Some surprising findings on the involvement of the parietal lobe in human memory. Neurobiol Learn Mem 2009;91:155–165.

43. Wager T, Smith E. Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci 2003;3:255–274.

44. Hoekzema E, Carmona S, Tremols V, et al. Enhanced neural activity in frontal and cerebellar circuits after cognitive training in children with attention-deficit/hyperactivity disorder. Hum Brain Map 2010;31:1942–1950.

45. Volkow ND, Wang GJ, Newcorn J, et al. Depressed dopamine activity in caudate and preliminary evidence of limbic involvement in adults with attention-deficit/hyperactivity disorder. Arch Gen Psychiatry 2007;64:932–940.

46. Milich R, Loney J, Landau S. Independent dimensions of hyperactivity and aggression: a validation with playroom observation data. J Abnorm Psychol 1982;91:183–198.

47. Milich R, Loney J, Roberts MA. Playroom observations of activity level and sustained attention: two-year stability. J Consult Clin Psychol 1986;54:272–274.

48. Roberts MA. A behavioral observation method for differentiating hyperactive and aggressive boys. J Abnorm Child Psychol 1990;18:131–142.

49. Loney J, Carlson GA, Salisbury H, Volpe RJ. Validation of three dimensions of childhood psychopathology in young clinic-referred boys. J Atten Disord 2005;8:169–181.

50. Breen MJ. Cognitive and behavioral differences in AdHD boys and girls. J Child Psychol Psychiatry 1989;30:711–716.

51. Schweitzer JB, Sulzer-Azaroff B. Self-control in boys with attention deficit hyperactivity disorder: effects of added stimulation and time. J Child Psychol Psychiatry 1995;36:671–686.

52. Breslau J, Miller E, Breslau N, Bohnert K, Lucia V, Schweitzer J. The impact of early behavior disturbances on academic achievement in high school. Pediatrics 2009;123:1472–1476.

53. Raggi VL, Chronis AM. Interventions to address the academic impairment of children and adolescents with ADHD. Clin Child Fam Psychol Rev 2006;9:85–111.

54. Hinshaw SP. Academic underachievement, attention deficits, and aggression: comorbidity and implications for intervention. J Consult Clin Psychol 1992;60:893–903.

55. Barkley RA, Fischer M, Smallish L, Fletcher K. Young adult outcome of hyperactive children: adaptive functioning in major life activities. J Am Acad Child and Adolesc Psychiatry 2006;45:192–202.

56. Polderman TJ, Boomsma DI, Bartels M, Verhulst FC, Huizink AC. A systematic review of prospective studies on attention problems and academic achievement. Acta psychiatrica Scandinavica 2010;122:271–284.

57. Breslau J, Miller E, Joanie Chung WJ, Schweitzer JB. Childhood and adolescent onset psychiatric disorders, substance use, and failure to graduate high school on time. J Psychiatr Res 2011;45:295–301.

58. World Medical Association. Declaration of Helsinki, Finland, 2008.

59. Barkley RA. Attention deficit hyperactivity disorder: a handbook for diagnosis and treatment. New York: The Guilford Press, 1990.

60. Barkley RA, McMurray MB, Edelbrock CS, Robbins K. The response of aggressive and nonaggressive ADHD children to two doses of methylphenidate. J Am Acad Child Adolesc Psychiatry 1989;28:873–881.

61. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986;42:121–130.

62. Volkow ND, Wang GJ, Fowler JS, et al. Dopamine transporter occupancies in the human brain induced by therapeutic doses of oral methylphenidate. Am J Psychiatry 1998;155:1325–1331.

63. Volkow ND, Wang GJ, Kollins SH, et al. Evaluating dopamine reward pathway in ADHD: clinical implications. JAMA 2009;302:1084–1091.

64. Johansen EB, Killeen PR, Russell VA, et al. Origins of altered reinforcement effects in ADHD. Behav Brain Funct 2009;5:7.

65. Waltz JA, Schweitzer JB, Gold JM, et al. Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. Neuropsychopharmacology 2009;34:1567–1577.

66. Swanson J. Compliance with stimulants for attention-deficit/hyperactivity disorder: issues and approaches for improvement. CNS Drugs 2003;17:117–131.

67. Weiss MD, Gadow K, Wasdell MB. Effectiveness outcomes in attention-deficit/hyperactivity disorder. J Clin Psychiatry 2006;67 (suppl 8):38–45.