Psychological Medicine

cambridge.org/psm

An integrated program of computer-presented and physical cognitive training exercises for children with attention-deficit/hyperactivity disorder

Bruce E. Wexler[1], Lawrence A. Vitulano[2], Christina Moore[2,3], Liliya Katsovich[2], Stephanie D. Smith[2,4], Cindy Rush[5], Heidi Grantz[2], Jinxia Dong[6] and James F. Leckman[2]

[1]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA; [2]Child Study Center, Yale School of Medicine, New Haven, CT, USA; [3]Department of Psychology, University of Delaware, Newark, DE, USA; [4]Department of Psychology, University of Southern Mississippi, Hattiesburg, MS, USA; [5]Department of Statistics, Columbia University, New York, NY, USA and [6]Peking University, Beijing, P.R. China

Original Article

Cite this article: Wexler BE, Vitulano LA, Moore C, Katsovich L, Smith SD, Rush C, Grantz H, Dong J, Leckman JF (2021). An integrated program of computer-presented and physical cognitive training exercises for children with attention-deficit/hyperactivity disorder. *Psychological Medicine* **51**, 1524–1535. https://doi.org/10.1017/S0033291720000288

Received: 10 August 2019
Revised: 21 January 2020
Accepted: 28 January 2020
First published online: 24 February 2020

Key words:
attention deficit hyperactivity disorder; cognitive training; non-pharmacologic treatment; physical exercise

Author for correspondence:
Bruce E. Wexler,
E-mail: bruce.wexler@yale.edu

Abstract

Background. This study integrated an experimental medicine approach and a randomized cross-over clinical trial design following CONSORT recommendations to evaluate a cognitive training (CT) intervention for attention deficit hyperactivity disorder (ADHD). The experimental medicine approach was adopted because of documented pathophysiological heterogeneity within the diagnosis of ADHD. The cross-over design was adopted to provide the intervention for all participants and make maximum use of data.

Methods. Children ($n = 93$, mean age 7.3 +/− 1.1 years) with or sub-threshold for ADHD were randomly assigned to CT exercises over 15 weeks, before or after 15 weeks of treatment-as-usual (TAU). Fifteen dropped out of the CT/TAU group and 12 out of the TAU/CT group, leaving 66 for cross-over analysis. Seven in the CT/TAU group completed CT before dropping out making 73 available for experimental medicine analyses. Attention, response inhibition, and working memory were assessed before and after CT and TAU.

Results. Children were more likely to improve with CT than TAU (27/66 *v.* 13/66, McNemar $p = 0.02$). Consistent with the experimental medicine hypotheses, responders improved on all tests of executive function ($p = 0.009$–0.01) while non-responders improved on none ($p = 0.27$–0.81). The degree of clinical improvement was predicted by baseline and change scores in focused attention and working memory ($p = 0.008$). The response rate was higher in inattentive and combined subtypes than hyperactive-impulsive subtype ($p = 0.003$).

Conclusions. Targeting cognitive dysfunction decreases clinical symptoms in proportion to improvement in cognition. Inattentive and combined subtypes were more likely to respond, consistent with targeted pathology and clinically relevant heterogeneity within ADHD.

Introduction

Attention deficit hyperactivity disorder (ADHD) is a neurodevelopmental disorder affecting 5–6% of school-aged children, disrupting executive cognitive functions (Barkley, 1997) and often leading to a variety of academic problems and social difficulties (Abikoff, Gittelman-Klein, & Klein, 1977; Barkley, Anastopoulos, Guevremont, & Fletcher, 1992; Barkley, DuPaul, & McMurray, 1990; Erhardt & Hinshaw, 1994; Whalen, Henker, & Dotemoto, 1980) and later maladaptive life outcomes (Cherkasova, Sulla, Dalena, Pondé, & Hechtman, 2013; Dalsgaard, Østergaard, Leckman, Mortensen, & Pedersen, 2015; Faraone, Sergeant, Gillberg, & Biederman, 2003; Huizink, Van Lier, & Crijnen, 2008; Popper, 1988; Wilens et al., 2002). Like most symptom-based psychiatric diagnostic categories (e.g. Hyman, 2010; Miller, 2010; Wexler, 1992), ADHD appears to include individuals who differ in underlying pathophysiology and etiology. As Coghill, Seth, and Mathews (2014) write, ADHD 'is an exemplar of a robust clinical neuropsychiatric syndrome with marked heterogeneity across multiple levels of analysis.'

For example, multiple neuropsychological tests differentiate groups of children with ADHD from typically developing (TD) children but each shows abnormality in only a minority of ADHD children (Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005), and studies employing multiple measures reveal independent dimensions of neuropsychological dysfunction that characterize subgroups of ADHD children (e.g. Coghill *et al.*, 2014; Sonuga-Barke, Bitsakou, & Thompson, 2010). Parent ratings of temperament have also identified distinct subgroups of ADHD, which were then validated by differences among the groups in cardiac physiology,

© The Author(s), 2020. Published by Cambridge University Press

CAMBRIDGE UNIVERSITY PRESS

resting state connectivity on functional magnetic resonance imaging (fMRI), and clinical course (Karalunas et al., 2014). Latent class analysis of personality traits identified subgroups of ADHD that differed in proportions of ADHD subtype or psychiatric co-morbidities (Martel, Von Eye, & Nigg, 2010). Electroencephalography power spectrum analyses identified subgroups that also differ in clinical symptoms and co-morbidities (Clarke et al., 2011), and have been replicated in different labs. Genetic and brain imaging studies add further evidence of the heterogeneity by showing multiple abnormalities but inability to establish 'signature' abnormalities for the clinical symptom-based diagnosis as an entity. Nearly 3000 genetic studies, including 32 meta-analyses, demonstrate strong heritability but no defining features of ADHD *per se* (Schachar, 2014). fMRI studies have led to a similar picture. The ADHD-200 Consortium (2012) challenged research teams to differentiate children with ADHD from TD children in a data set of fMRI data, symptom ratings, age, gender, handedness, and medication history. Of the children identified as ADHD by the top performing team using the fMRI data, 94% were in fact ADHD. However, they only identified 21% of the ADHD cases. The team that was most accurate overall made predictions only on demographic data.

One way to develop new and potentially specific treatments in the presence of such pathophysiological and probable etiological heterogeneity within a diagnostic category is to develop a treatment that targets a core aspect of pathology present in some individuals with the diagnosis. While such a treatment would not be of benefit for individuals in whom that particular pathophysiology was not relevant to their condition, it could be of considerable benefit for those individuals in whom the targeted dysfunctions are central to their illness. This approach has been codified in the steps that define the experimental medicine approach described by the Science of Behavioral Change Working Group (Riddle & Science of Behavior Change Working Group, 2015) and endorsed by the National Institute of Mental Health: (1) identify an intervention target, (2) develop measures to verify the target, (3) engage the target through intervention, and (4) test the degree to which target engagement produces desired behavior change.

In previous work in general populations of children and in studies of children with ADHD, we and others have addressed several steps of the experimental medicine approach described by the Science of Behavioral Change Working Group (Riddle & Science of Behavior Change Working Group, 2015). In a study of over 1000 elementary school children we demonstrated ability to measure and improve the executive function targets (steps 2 and 3; Kavanaugh, Tuncer, & Wexler, 2019). Then specifically in ADHD, we showed that the intervention impacts attention-related ERPs during the Go/No-Go task (step 3; Smith et al., 2019), and another group showed effects of the intervention on regional brain activations during attention and working memory tasks (also step 3; Rosa et al., 2019). The present report shows that clinical improvement is related to the degree to which the intervention engages three executive function targets (step 4).

In the present study we targeted and assessed three executive cognitive dysfunctions previously shown to be compromised in significant numbers of children with ADHD: focused attention in the presence of distraction; response inhibition or self-control; and working memory. While the selection of target cognitive dysfunctions is complicated by the very heterogeneity we attempt to address, these dimensions of cognitive dysfunctions have been shown in repeated studies to differentiate children with ADHD from TD children with effect sizes similar to or higher than found in other dimensions of cognitive dysfunction (e.g. Coghill et al. 2014; Fair et al., 2012; Fried, Hirshfeld-Baker, Petty, Batchelder, & Biederman, 2015; Wahkstedt, Thorell, & Bohlin, 2008). The intervention consisted of three components: (1) computer-presented executive-function cognitive training (CT) exercises; (2) cognition-enriched physical exercises designed to require executive cognitive functions in the context of whole-body activity and social interaction; and (3) a group-level game shown to increase participation in class activities.

While past work has shown that computer-presented cognitive-training exercises can reduce clinical symptoms, normalize task-related regional brain function, improve cognition, and lead to better real-world function in patients with schizophrenia and depression (e.g. Morimoto et al., 2014; Wexler & Bell, 2005; Wykes, Huddy, Cellard, McGurk, & Czobor, 2011), meta-analyses show much more limited effects in children with ADHD (Cortese et al., 2015; Melby-Lervåg & Hulme, 2013; Rapport, Orban, Kofler, & Friedman, 2013; Sonuga-Barke et al., 2013). In general, these interventions have led to improvement in the cognitive function that was specifically trained (i.e. attention or spatial memory) but evidence of generalization to other aspects of compromised cognition was limited and ADHD clinical symptom reductions were inconsistently seen (Beck, Hanson, Puffenberger, Benninger, & Benninger, 2010; Chacko et al., 2014; Cortese et al., 2015; Gray et al., 2012; Green et al., 2012; Holmes et al., 2010; Johnstone et al., 2012; Klingberg et al., 2005; Shalev, Tsal, & Mevorach, 2007; van Dongen-Boomsma, Vollebregt, Slaats-Willemse, & Buitelaar, 2013). There are, however, also some studies that support the potential value of CT for children with ADHD. Holmes et al. (2010) found that 20–25 sessions of computerized spatial memory training over 4 weeks led to significant gains in spatial working memory and also in verbal and auditory working memory measured by tasks very different in content and format from the training exercises. The gains were above the effects of medication and maintained 6 months after training. In a random-assignment wait-list-controlled study of combined spatial memory and inhibitory control training in 60 children with ADHD, Johnstone et al. (2012) found significant symptom reduction and increased performance on focused attention, target detection, and auditory working memory tasks, all different from the training tasks. Dovis, Van der Oord, Wiers, and Prins (2015) further expanded the range of executive function training to include working memory, inhibition and cognitive flexibility, and randomly assigned children to 25–30-min sessions of either training or a control condition of similar tasks without key types of training trials and fixed at minimal difficulty levels. Treatment gains were significant in an omnibus evaluation of six different measures of executive function, with effect sizes medium to large on three of the tests. The present study further expanded the computer-presented exercises to train sustained and focused attention, response inhibition, cognitive flexibility, working memory, pattern recognition, and category formation. In addition, the computer programs included newly developed algorithms that adapt the training to the individual user more powerfully than has previously been possible.

An earlier paper reported results of an intention-to-treat parallel group analysis that compared only the Fall groups (i.e. treatment *v.* control groups before the cross-over) and only during the first 2 years of the present study. Effects of treatment on symptom ratings by clinical assessors, teachers or parents were not statistically significant in that sample (Smith et al., 2016). In contrast, the present paper reports clinical and cognitive outcomes in the full

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

sample of children who received the intervention in either the Fall or Spring and includes children from the third and final year of the study thereby increasing the sample of children who received the program to 73 from 42 in the initial report by Smith et al. (2016). In addition, the present report gains the added power of the cross-over analysis (Dwan, Tianjing, Altman, & Elbourne, 2019), and follows the experimental medicine approach to deal with the pathophysiological heterogeneity in the study population and evaluate the relations among CT, cognitive improvement, and symptom reduction.

Methods

Participants

Participants were recruited via letters describing the study to parents of children in kindergarten, first, and second grades in participating schools. The SNAP-IV rating scale (Swanson, 1992) assessing ADHD symptomatology accompanied this letter and if the parent/guardian completed the rating scale, the SNAP-IV was also completed by the child's teacher. All study procedures comply with ethical standards of the Human Investigation Committee at the Yale School of Medicine and with the Helsinki Declaration of 1975, as revised in 2008. Informed consent was obtained from parents or guardians, and all participants gave informed assent prior to engaging in any study procedures. Families were paid $40 for their time following each assessment visit.

A child was considered 'screen-positive' if the average rating per item was greater than 1.2 on the parent or teacher SNAP-IV (Bussing et al., 2008), and invited to participate in a baseline assessment using the Kiddie Schedule for Affective Disorders and Schizophrenia – Present and Lifetime Version (K-SADS-PL), administered by doctorate or master-level clinicians as an interview with the child's parent/guardian. During baseline evaluations, the parent SNAP-IV was completed for a second time facilitated by research staff. Inclusion criteria were: (a) age 5 to 9 years; (b) diagnosis of ADHD according to the DSM-IV TR criteria for ADHD or children deemed to be at 'high-risk' for ADHD defined as one symptom below diagnostic criteria; (c) an intellectual quotient of at least 80 on the Kaufman Brief Intelligence Test (K-BIT 2; Kaufman and Kaufman, 2004b); and (d) on a stable dose of medication for at least 4 weeks (if on medication for ADHD). Children were excluded if they had a co-morbid psychiatric diagnosis, acute behavior problems or physical disability that would prevent them from participating in treatment. A comorbid diagnosis of autism spectrum disorder was not itself exclusionary. Children were assessed before and after a 15-week period of participation in the CT program and a 15-week period of treatment as usual (TAU) both within the same school year. Only about 15% of the children were receiving medication. All children were in mainstream classrooms and while teacher ratings indicated they were aware of attention and self-control problems in many of the participants, any non-study interventions were almost entirely limited to individual teacher classroom management strategies and there was no indication that these differed during the intervention and TAU periods. Forty-eight children were randomly assigned to receive the CT in the first 15-week period and 45 children in the second 15-week period. Of the 93 enrolled, 15 dropped out of the CT/TAU sequence group and 12 out of the TAU/CT sequence group before completing both periods, leaving 66 available for the cross-over analysis. Seven in the CT/TAU group completed CT before dropping out to make a total of 73 available for the experimental medicine analyses (see Fig. 1 for the consort flow diagram, Table 1 for subject characteristics). Regarding specifically dropouts during the CT, 13 of the 86 children (15%) who began a 15-week intervention period failed to complete pre- and post-intervention assessments. Two did not begin the intervention. Parents indicated that the 11 who dropped out after starting the program did so for reasons ranging from not liking the program to not having enough time for homework.

The integrated three-part intervention

Computer-presented cognitive-training exercises

The CT games were designed by BEW and developed and supported as web-based applications by the Yale startup company C8Sciences. The original version consisted of three games, each with 80–150 levels of difficulty. The first game begins with the child having to click on a yellow ball moving randomly across the screen whenever it turns red, exercising sustained attention. The ball moves faster following correct responses and slows after errors. As they either reach a preset high level of performance or stay at a lower performance level without improvement for an extended period of time, the child is moved through progressive levels that layer in additional cognitive demands. On the next level the ball sometimes turns blue (a foil) that is to be ignored, adding response inhibition. Next, the target color randomly changes back and forth between blue and red, increasing required response inhibition and adding cognitive flexibility. Next levels require working memory as a ball is a target only if it changes to the same color twice in a row ('same as last') or changes to a different color ('different from last'). All rules are repeated with two and then three balls on the screen. In the second game, children click on butterflies carrying signs only if the object on the sign was a member of a designated category (e.g. animals. furniture, tools, machines). With correct responses, the butterflies move faster and more butterflies are on the screen at the same time (from 1 to 6). At higher levels, categories rotate, two categories are targets simultaneously, or the child must find two objects on the screen in the same category. The third game requires the child to figure out the rule that links a series of three objects and use this rule to choose the fourth object to complete the row. Time to respond becomes shorter with correct responses and rules become more complex in higher levels. Points earned for correct responses were used each day to get rewards from a virtual store. The artwork for the game-play screens was changed in year three of the study to increase student engagement, however, the underlying computer code and CT progression of the games remained unchanged. In addition, a simple spatial span game that changed only in list length was added as a repeated assessment and limited memory training activity for a small proportion of game play time (<20%). Given that the neural-system dysfunction targeted by the overall intervention was unchanged, and the outcomes of interest were generalized clinical and cognitive benefit rather than improvements in the specific CT exercises themselves, data from all 3 years were used in the analyses. Outcomes in the third year did not differ from those in the first 2 years.

Physical exercises

Like the computer exercises, cognitive aspects of the physical exercises begin with sustained attention and response inhibition,

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

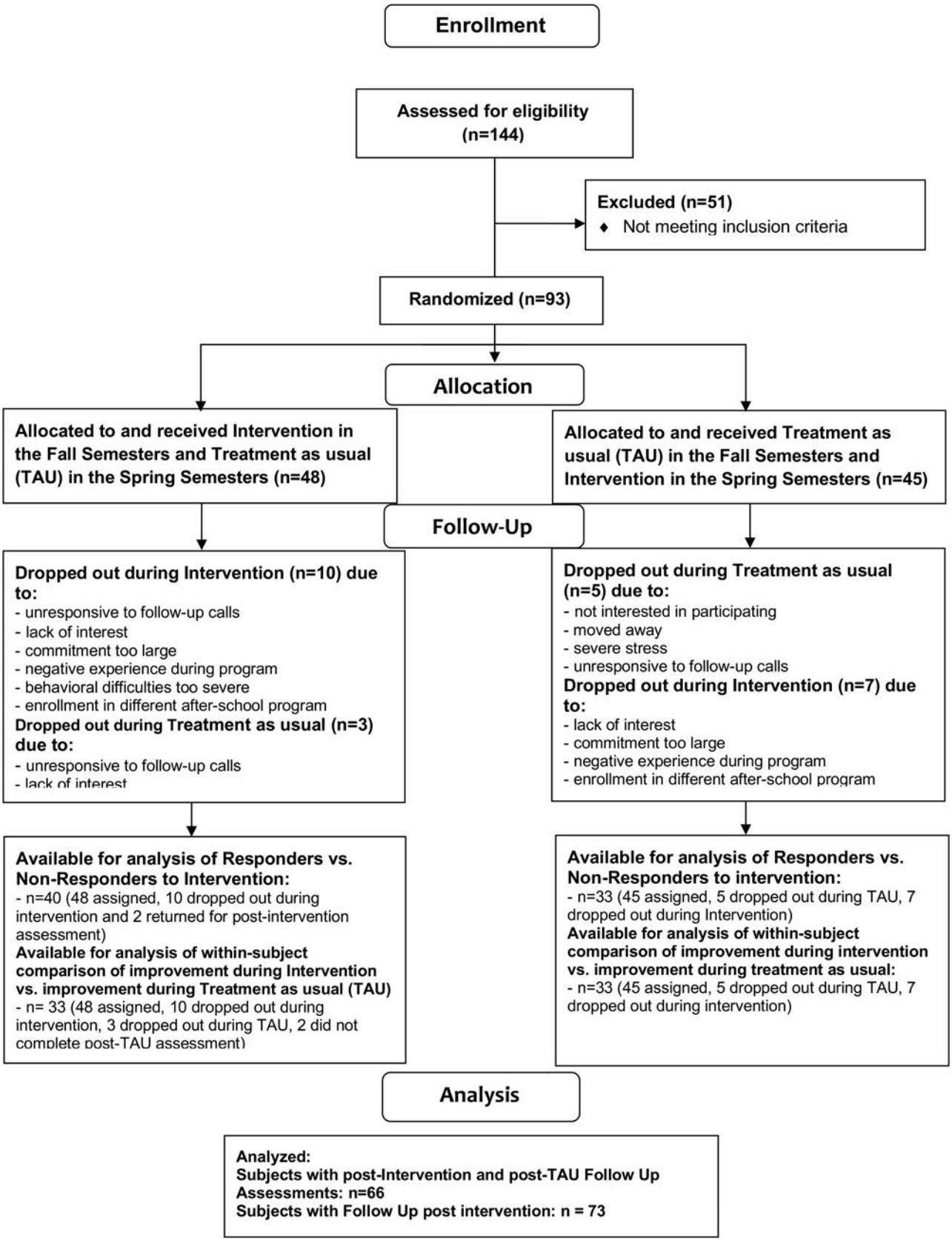


Fig. 1. Consort diagram.

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

Table 1. Demographic and clinical characteristics

	All subjects *n* = 73	Inattentive/ combined subtype *n* = 55	Hyperactive subtype *n* = 18
Age, mean (s.d.) in years	7.4 (1.2)	7.3 (1.2)	7.5 (1.2)
Sex, No. (%)			
Male	48 (66)	34 (62)	14 (78)
Female	25 (34)	21 (38)	4 (22)
Race, No. (%)			
White	31 (43)	19 (35)	12 (67)
African American	17 (23)	15 (27)	2 (11)
Hispanic	14 (19)	11 (20)	3 (17)
Asian	1 (1)	1 (2)	–
Other	10 (14)	9 (16)	1 (5)
ADHD diagnosis total, No. (%)	62	49	13
Inattentive	23 (31)	23 (42)	–
Hyperactive	13 (18)	–	13 (72)
Combined	26 (36)	26 (47)	–
At risk for ADHD total, No. (%)	11	6	5
Inattentive	6 (8)	6 (11)	-
Hyperactive	5 (7)	–	5 (28)
Combined	–	–	–
Parental education, years			
Mean (s.d.)	15.1 (2.5)	14.8 (2.6)	16.1 (2.1)
Comorbidity, No. (%)			
Tic disorder	1 (1)	1 (2)	–
Depression	2 (3)	2 (4)	–
Anxiety disorder	16 (22)	12 (22)	4 (22)
Enuresis	6 (8)	4 (7)	2 (11)
ODD	14 (19)	9 (16)	5 (28)
ASD	3 (4)	2 (4)	1 (6)
Speech problems	3 (4)	3 (5)	–

No statistically significantly differences were found between inattentive/combined subtype *v.* hyperactive subtype.

and progressively layer in cognitive flexibility, multiple simultaneous attention, and working memory. For example, initially children are each assigned their own space within their own circle on the floor, attend to their own bodies and practice yoga-like balancing poses. Next they do controlled ball passing in pairs, group running games with rules that require planning, strategy and self-control, or response inhibition games like 'Simon Says.' Later they learn martial arts and dance sequences, or throw two different colored bags to one another in circles of 5–6 children, with each color having a different sequence of individuals to whom it is thrown. Each day there is a mix of more and less aerobic games, and group and individual focused exercises.

Social component GBG

Children were divided into two teams that could earn rewards by following the rules that encouraged behaviors important in creating classroom learning environments (e.g. waiting your turn, staying in your seat, Good Behavior Game (GBG), Embry, 2002). Each team could win by having few enough rule violations. 'Wacky Behavior' rewards (e.g. 'follow the leader funny walking') were given to winning teams immediately following a game, and supplemented by prizes at the end of each week for the team with the most wins.

Procedure

Training and implementation

Children participated after school, 3–4 times per week with 45 min in the computer lab (30 min of actual computer exercises) and 45 min of physical exercises. Teachers were trained and paid to implement the program with assistants provided by the research team. In order that participation in the program not identify a child as having ADHD, and help create a growth-enhancing environment, approximately half the children in each class were TD children who volunteered for the program as a free after school enrichment program and were not research subjects. There were 6–10 children in each group. Continuous data capture and pre-determined schedule of computer games supported implementation fidelity for this component of treatment. Staff were trained in the physical exercise program and GBG with additional instruction throughout the program. Research staff used observational checklists to confirm fidelity of the GBG, which appeared adequate. However, variation in experience with physical education among teachers and the inability to monitor and provide feedback to individual children limited the consistency of the physical exercise component.

Behavioral measures

Parent SNAP-IV ratings completed before and after the intervention and TAU periods served as primary ratings. Parent ratings have been most sensitive to change in previous studies of ADHD (e.g. Beck et al., 2010; Chacko et al., 2013; Cortese et al., 2015; Klingberg et al., 2005) perhaps because parents see children across settings and ADHD symptom manifestation is sensitive to context (e.g. Barkley, 2003; Cortese et al., 2015). Teacher SNAP-IV ratings (when available) and clinician SNAP-IV ratings based on a review of parent and teacher ratings, and a clinical interview with parents, were collected at the same times.

Cognitive assessment measures

Three dimensions of the executive cognitive function shown to be compromised in previous studies of children with ADHD and identified as intervention targets – focused attention, response-inhibition, and working memory – were assessed with web-based measures automatically presented, administered, and scored in the classroom setting of the brain-training program itself. Given that ADHD symptoms manifest differently in different contexts, we thought it was important to administer assessments in a real-world environment that would require functional demands of the child similar to those in everyday life. The three tests were given one per day beginning on the third day of the program and again one per day in the next to last week of the program. Two tests followed precisely the design of tests in the NIH Toolbox of tests of executive function (nihtoolbox.org). The first was the Flanker Test of focused attention where the primary performance measures

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

were percent correct and reaction time on correct incongruent trials. In this task, children have to indicate by keyboard response the pointing direction (right or left) of the center arrow in a linear horizontal array of five arrows. On incongruent trials, the four 'flanking' arrows point in the opposite direction of the central arrow. Scores of less than 65% correct on the congruent trials suggest that the child did not understand or engage with the test, and these children were not considered in the analyses. The second test was the List Sorting Working Memory Test. Subjects are shown a series of animals or household objects. They then have to click on the objects they have just seen in a grid of 16 objects in order from smallest to largest rather than the order in which they were presented. The test starts with a list of 2 objects. If the subject completes the list accurately, list length is increased by one. If they err, the same length list is repeated. Two failed attempts at the same list length end the test. The score is the sum of correct list lengths. In part one, trials of animals and household objects alternate. In part two, animals and household objects are presented in the same trial, and subjects have to reorder the animals first and then the household objects. Children who scored 0 on part one of the tests were not included in the analyses. The third test is a Go/No-Go test of response inhibition. Subjects are instructed to press the space bar whenever a 'Go' stimulus is presented but not when a 'No-Go' stimulus is presented. There are three blocks with different stimuli, 50 stimuli per block with 40 Go and 10 No-Go trials, randomized in sets of 10 with 8 Go and 2 No-Go in each set. In the first block 'P' is the go stimulus and 'R' is the no-go stimulus. In the second block this is reversed. In the third block, pictures of furniture are go trials and pictures of foods like cake and ice cream are no-go stimuli. Stimuli are presented for 400 ms with a 1400 ms response window after stimulus offset. Errors are indicated by the display of a large red 'X.' Children who scored less than 90% on the Go-Trials were not considered in the analyses since failure to respond rates greater than 10% can artificially inflate No-Go accuracy. Absences from the program on a day that a pre- or post-test was administered meant that data on that test was unavailable for the child.

Statistical analysis

The first research question was: 'Are children more likely to show substantial reductions in symptoms after participation in the program than during treatment as usual?' Substantial reduction of symptoms (i.e. 'Improver' or 'Responder' status) was defined as a 30% reduction in parent ratings of overall symptoms. This threshold was based both on the levels of improvement in parent ratings reported in past studies following pharmacologic or behavioral treatments. The MTA study found ~35% mean reduction in parent ratings with medical treatment and 20% with behavioral treatment (Jensen, 2003). Subsequent studies of methylphenidate (Abikoff et al., 2004) and atomoxetine (Allen et al., 2005) reported mean reductions in parent ratings of 36.8% and 28% respectively. Previous studies with another CT program, CogMed, reported reductions of parent ratings of inattentive symptoms of 30% and hyperactive symptoms of 25% (Klingberg et al., 2005), and 11% in the overall Conners' ADHD index (Beck et al., 2010). Thirty percent is also consistent with criteria used in studies to define the treatment response in other disorders (e.g. Bloch et al., 2006; Pallanti, Querciol, Sood, & Hollander, 2002; Storch, Lewin, De Nadai, & Murphy, 2010; Suzuki et al., 2012). A symptom reduction threshold of 30%, then, requires that a child improve as much as the average child improved in multiple past studies showing significant treatment effects. Consistent with CONSORT recommendations for cross-over designs, the McNemar test was used to compare proportions of children showing 30% reductions in parent ratings of symptoms during the intervention and TAU periods (Dwan et al., 2019).

Most important, however, in the analysis and this report, is the second research question: 'Is clinical response related to change in the cognitive functions the intervention is designed to address?' The first analysis related to this question determined whether children whose parents report at least a 30% reduction in symptoms after doing the program ('responders') showed greater improvement in cognition than 'non-responders'? An affirmative answer would demonstrate the hypothesized association between improvement in the cognitive functions targeted by the intervention and symptom improvement. Moreover, if groups that were defined on the basis of parent assessments of symptom reduction also differed in the amount of improvement on objective tests of cognition, it would provide confirmation that the parents were correct in their assessments of whether their child had responded to the intervention. We analyzed between-groups effects for responder and non-responder status, within-groups main effects of time, group-by-time interactions, and improvement for responders and non-responders separately. A second analysis treated improvement in parent ratings as a continuous rather than dichotomous variable, and used a linear model to predict parent SNAP percent change. Finding that changes in cognition predict clinical response to the intervention would support the experimental medicine study hypothesis that targeting cognitive dysfunction in children with ADHD is an effective way to treat the disorder in some children, and would further support the meaningfulness of the parent ratings. To build the model we considered pre-test and difference scores for the following measures as predictors: Working Memory total score, Flanker Incongruent trial accuracy, Flanker Incongruent trial reaction time (on correct trials), and GNG No-Go trial accuracy. In obtaining the final model, we began with a full model including the pre-test and change scores for each of the four measures, and sequentially eliminating the least significant predictors until only significant ($p < 0.10$) predictors remained and removal of the least significant of these markedly degraded the model (indicating that it captured substantial variance independent of the other remaining predictors). Students were eliminated who were missing all three tests at baseline or post-intervention ($n = 10$) and missing data from other students was imputed using average scores (for each test/measure).

Even though hypotheses for questions one and two were directional, all p values reported are two-tailed (including for the interaction terms). Given the small number of a priori comparisons, no further corrections were made for multiple comparisons.

All calculations to address the relationship between target impact (i.e. cognitive function) and symptom reduction were done in R using the RStudio interface running R version '3.1.2 (31 October 2014)' and RStudio version 0.98.1091. None of the tests or calculations used Flanker scores for students with less than 65% accuracy on Flanker Congruent Trials (this removes the Flanker test scores of five students). None of the tests or calculations used GNG scores for students with less than 90% accuracy on GNG Go Trials (this removed the GNG test scores of 19 students). This analysis also removes the Working Memory scores of one student whose values were multiple standard deviations above the rest and the Flanker scores of two students whose

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

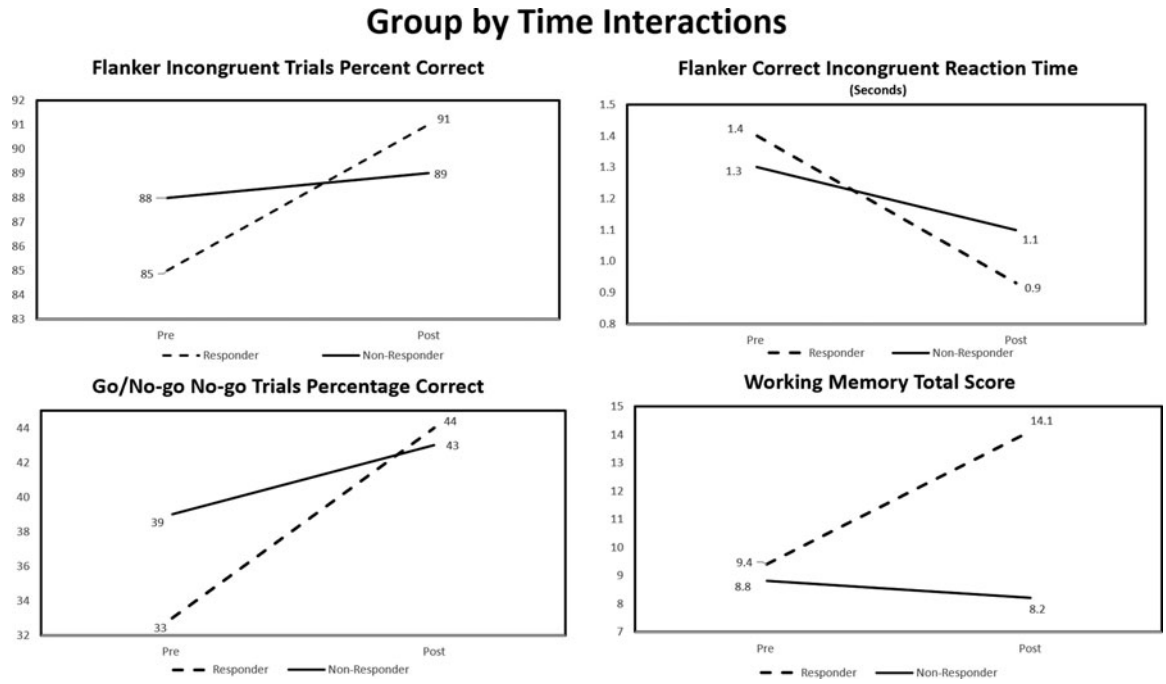


Fig. 2. Changes in Web-Based Classroom Administered Tests of Executive Function in responders and non-responders.

reaction times scores indicated mistrials. The linear model was built using the lm function in R on the eight predictor values for all students in the dataset.

The above analyses constitute the proof tests of study hypotheses. Given that these primary analyses supported the existence of subgroups characterized by response to this type of CT intervention, secondary analyses compared responders and non-responders with regard to demographic and clinical features for the dual purposes of providing further validation and characterization of these subgroups. Means and standard deviations for all clinical ratings pre and post intervention or TAU periods for all 66 children are presented as supplementary material.

Results

Research question one

'Are children more likely to show substantial reductions in symptoms after participation in the program than during treatment as usual?'

In the active treatment period, 27 of 66 children (41%) showed at least a 30% reduction in parent ratings of ADHD symptoms while during TAU only 13 of the 66 children (20%) did. The greater proportion of 'improvers' following active treatment period was significant by McNemar's test, χ (1) = 5.76, $p = 0.02$. Results are essentially no different when considering only the 62 children who did at least 20 sessions of the computer and physical exercises. None of the four children who participated in less than 20 sessions had 30% symptom reduction. When only children with definite ADHD are considered the response rate is 43.5%. When children who dropped out after beginning the intervention are considered treatment failures, the response rate is 38%.

The intervention and TAU groups had nearly identical parent ratings of ADHD symptoms at the study entry (intervention 28.1 +/− 11, TAU 28.2 +/− 10.4, $p = 0.97$) and after the cross-over (intervention 24.5 +/− 7.2, TAU 25.2 +/− 10.6, $p = 0.77$), indicating that differences in baseline scores between groups either before or after cross-over did not affect the analysis. The ratio of symptom reduction during CT to TAU (CT/TAU) was nearly identical in the CT/TAU (1.53) and the TAU/CT (1.51) sequence groups, indicating that order of the cross-over sequence did not affect the results. The number of dropouts was also independent of sequence with 13 dropouts in the group that began with the intervention and 12 in the group that began with TAU.

Research question two

'Is clinical response related to change in the cognitive functions the intervention is designed to address?'

Responders showed greater improvement than non-responders on all three cognitive tests (Fig. 2, Table 2). The group-by-time interaction was significant for performance on the working memory test ($F_{(1,40)} = 4.83$, $p = 0.03$), approached significance for accuracy on incongruent trials in the Flanker test ($F_{(1,51)} = 3.7$, $p = 0.06$) and reaction time on correct incongruent Flanker trials ($F_{(1,51)} = 2.8$, $p = 0.09$), and was not significant for No-Go Trial accuracy ($F_{(1,38)} = 1.60$, $p = 0.21$). *Post-hoc* tests revealed significant improvement on each test only in the responders: working memory performance responders pre-mean 9.4/post-mean 14.1 $p = 0.009$, non-responders 8.8/8.2 $p = 0.75$; Flanker incongruent accuracy responders pre/post 85%/91% $p = 0.02$, non-responders 89%/89% $p = 0.81$; and reaction time on correct incongruent Flanker trials responders 1.4/0.9 s $p = 0.003$, non-responders 1.3/1.1 s $p = 0.27$. No-Go Trial accuracy followed a similar pattern although the group-by-time interaction did not approach significance: responders 33%/44% $p = 0.01$, non-responders 39%/43% $p = 0.39$. Baseline scores of responders and non-responders did

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

Table 2. Neurocognitive tests in responders and non-responders pre/post intervention

	Responders ($N = 31$)			Non-responders ($N = 42$)			Interaction	
	Baseline	Post treatment	p value	Baseline	Post treatment	p value	p value	Effect size
Working memory test	9.4 (4.5)	14.1 (8.2)	0.009	8.8	8.2	0.75	0.03	0.65
Flanker accuracy incongruent trials	85% (16)	91% (11)	0.02	89% (12)	89% (11)	0.81	0.06	0.52
Flanker reaction time	1.4 (0.9)	0.9 (0.3)	0.003	1.3 (0.6)	1.1 (0.7)	0.27	0.09	0.46
No-Go trial accuracy	33% (14)	44% (20)	0.01	39% (19)	43% (19)	0.39	0.21	0.40

Responders defined on the basis of at least 30% reduction in Parent SNAP ratings, showed a significant improvement on all four neurocognitive measures (0.02 to 0.009) while non-responders did not improve significantly on any measure (0.27 to 0.81).

Table 3. Regression model using baseline and improvement scores on tests of cognition to predict improvement in parent ratings of symptoms

Coefficient	Estimate	Std. error	t-Statistic	p value
Intercept	−37.14	17.64	−2.11	0.040
Difference flank CI	108.99	41.58	2.62	0.011
Pre flank RT	17.87	7.11	2.51	0.015
Pre WM	2.21	1.25	1.77	0.082
Difference WM	1.69	0.87	1.95	0.056

Regression coefficients and statistical significance of each in model relating neurocognitive baseline function and improvement to reduction in clinical symptoms. The overall model was significant at $p < 0.008$.

not differ significantly, nor did mean number of hours of computer training: responders 18.2 h (S.D. = 3.9, range = 8–25); non-responders 17.5 h (S.D. = 4.9, range = 4–25).

The linear regression model using baseline and change scores on the working memory, Flanker and GNG tests to predict the change in parent ratings as a continuous outcome measure was robust and highly significant at $p < 0.008$ ($F = 3.8$, df, 4, 58). Nothing in the model diagnostics leads to concerns about the assumptions of the model. Four potential predictors were dropped from the model (see methods), with minimal associated decreases in R^2 (0.220 to 0.219 to 0.219 to 0.215 to 0.209). Four predictors (Table 3) remained with omission of the least significant of the four dropping R^2 from 0.209 to 0.166.

Research question three

'Do "Responders" differ in baseline clinical features from "Non-Responders"?'

Diagnostic subtype

The treatment response varied significantly as a function of diagnostic subtype. For the inattentive and combined subtypes the treatment response rate was 49% (27 of 55 children) while for the hyperactive subtype it was only 22% (4 of 18), (χ^2 (1) = 4.0, $p = 0.045$). The difference is more pronounced when children with sub-threshold ADHD are excluded [53% (26/49) *v.* 8% (1/13), χ^2 (1) = 8.6, $p = 0.003$]. The average hours of computer training received by the hyperactive children (16.6, S.D. = 4.2) did not differ significantly from the other children (18.2, S.D. = 4.5). When children who dropped out after beginning the intervention are considered treatment failures, the response rate in inattentive or combined ADHD is 46%. At baseline, responders had significantly higher total parent SNAP scores [29.6 *v.* 24.5, t (71) = 2.27, $p = 0.03$] and higher inattentive subscale scores [16.0 *v.* 12.2, $t(71) = 2.72$, $p = 0.008$] subscale scores, but the groups did not differ significantly on the hyperactive subscale [13.6 *v.* 12.4, t (71) = 0.94, $p = 0.35$]. Interestingly, the responders significantly improved over time on hyperactive (13.6 to 7.1) as well as inattentive subscales (16.0 to 8.3), while non-responders improved in neither (hyperactive 12.4 to 13.5; inattentive 12.2 to 12.9), with the difference in improvement between groups highly significant for both subscales ($p<0.0001$).

Demographics

Gender, age, ethnicity, and K-BIT scores did not differ between responders and non-responders (Table 4). There was no difference in the proportion of children on medications for ADHD between responders and non-responders, 6 of 31 children (19%) and 7 of 42 (17%), respectively, or on stimulant medications in particular, 13% (4/31) in the responder group and 17% (7/42) in the non-responders group. Medication or dose changes were noted in 4 subjects among the responders and 2 subjects among non-responders.

Discussion

Children with ADHD or sub-threshold for ADHD participated in a program of computer-presented and physical exercises designed to improved cognitive functions commonly compromised in ADHD and thought to play a central role in generating clinical symptoms that define the disorder. Children whose parents indicated that their ADHD symptoms were reduced by at least 30% had highly significant gains in three 'gold standard' tests of executive function, while children whose parents indicated that their symptoms had not been reduced by at least 30% did not have significant gains on any of the tests. The degree of clinical symptom reduction in the entire sample was associated with the degree of improvement in the cognitive functions. These data suggest that there may be a subset of children with ADHD who favorably respond to a CT intervention with both reduced symptoms and improved cognitive functions thought to be central to the disorder. Treatment responders constituted 41% of the sample that remained available for assessments before and after both the intervention and TAU periods. When only children who meet full criteria for ADHD inattentive or combined subtype were considered, the response rate increased to 53%. Over 80% of those who met the 'responder' criterion were not taking medications, and the criterion for response, 30% reduction in symptoms, is close to the average improvement reported with pharmacotherapy. While a substantially higher proportion of children respond

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

Table 4. Responders *v.* non-responders

	Responders $n = 31$	Non-responders $n = 42$
Age, mean (s.d.) in years	7.2 (1.1)	7.6 (1.2)
Sex, No. (%)		
Male	18 (58)	30 (71)
Female	13 (42)	12 (29)
Race, No. (%)		
White	12 (39)	19 (45)
African American	8 (26)	9 (22)
Hispanic	6 (19)	8 (19)
Asian	–	1 (2)
Other	5 (16)	5 (12)
ADHD subtype, No. (%)		
Inattentive/combined	26 (84)	23 (55)
Hyperactive	1 (3)	12 (28)
At risk for ADHD	4 (13)	7 (17)
Medication for ADHD, No. (%)	6 (19)	7 (17)
Baseline Clinician SNAP, mean (s.d.)	27.9 (6.8)	26.7 (6.6)
Baseline Parent SNAP, mean (s.d.)*	29.6 (9.7)	24.5 (9.3)
IQ, mean (s.d.)	104.3 (11.9)	104.7 (14.1)
Computer Training Hours, mean (s.d.)	18.2 (3.9)	17.5 (4.9)

Responders had higher baseline Parent SNAP ratings *$p < 0.05$; there were no other significant differences between the group.

to stimulant medications (e.g. Barkley, 1977; Efron, Jarman, & Barker, 1997) than responded to the CT intervention, these data suggest that CT programs such as the one used in this study could be a useful treatment for some children with ADHD, before or after starting medications, with potentially fewer side effects than with medications.

Identification of possible pathophysiological processes through cognitive, neuroimaging, and neurochemical measures has created opportunities to target some of the putative pathophysiological processes and measure impact on both the pathology and clinical symptoms. Our intervention was designed to address compromised focused attention, response inhibition, and working memory, and positively impact the distributed neural functional systems associated with these executive cognitive operations. We found that improvement in accuracy on the Flanker Incongruent trials with distraction, and improvement in working memory both predicted improvement in parent ratings of clinical symptoms. This further supports the clinical value of CT programs that address these cognitive dysfunctions. The change in accuracy on the Flanker Incongruent trials (with distraction) and Working Memory showed very low correlation, indicating independent associations with clinical outcome. This might be an example of heterogeneity within the study population, with some children having a disorder that is marked primarily by easy distractibility and others by decreased memory. Alternately, it might be that the treatment is addressing similar neuropathology in most of the responders, but that the neuropathology has different cognitive manifestations in different children. We also found that the slower response when correctly identifying incongruent trials and higher working memory scores at the baseline predict the clinical response. The slower response on correct incongruent trials is understood as reflecting a greater effect of the distractors. This is consistent with the fact that the response rate was much higher among children meeting full criteria for ADHD inattentive or combined subtype (53%) than among those with Hyperactive subtype (8%). Together these observations support the common-sense expectation that children with difficulty focusing attention are most likely to benefit from our intervention. However, the fact that children with inattentive or combined subtype showed a significant improvement in hyperactive as well as inattentive symptoms suggests that the treatment is effective in particular types of children rather than for a particular type of symptom.

A recent fMRI study evaluated effects of the training program used in this study in 6–13 year-old children ($n = 10$) with ADHD who had persistent symptoms despite pharmacotherapy. When compared to an active control group of children with ADHD ($n = 10$) who received educational videos, after intervention the CT group showed greater activation increases in response to increased attention demands in bilateral precuneus, right insula, bilateral associative visual cortex, and angular gyrus, and right middle temporal, precentral, postcentral, superior frontal, and middle frontal gyri. During an N-back working memory task, after intervention the CT group showed smaller activation increases in response to increased load in right insula and putamen and left thalamus and pallidum (Rosa et al., 2019). While highly preliminary, these observations suggest that the multicomponent CT program used in this study engages and affects task-dependent neurocognitive systems that integrate activation across multiple brain regions. If these affects are activity-dependent neuroplastic changes, the affects could be more lasting than the effects of medications.

Parent ratings are often used in treatment studies because ADHD symptoms are context dependent and parents see their children for many hours and in varied settings. In our study and others (Amador-Campos, Forns-Santacana, Guàrdia-Olmos, & Peró-Cebollero, 2006; Fischer, Barkley, Fletcher, & Smallish, 1993; Sibley, Altszuler, Morrow, & Merrill, 2014) parent ratings do not correlate highly with teacher ratings, and in several past studies of CT interventions parent but not teacher ratings showed improvement (Beck et al., 2010; Klingberg et al., 2005). This discrepancy between the ratings of parents, who are emotionally involved with the child and not blind to treatment condition, and teachers, who in some ways are more objective, has raised questions of bias in parent ratings, thus potentially limiting their value in assessing treatment response. The present results provide external and objective evidence that the parent ratings accurately indicated whether their child had improved or not.

Limitations

Since the intervention has three different components it is not possible to know if all components are needed. The absence of an active control condition makes it impossible to know the degree to which factors such as the added structured activity after school contributed to benefits during the intervention period. However, this concern is mitigated by the fact that, following the experimental-medicine model, the greater the degree to which the intervention successfully impacted the executive cognitive function pathology it was designed to address, the greater the reduction in clinical symptoms. These predicted differences among subjects all of whom were in the same active treatment condition are not

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

impacted by differences between the intervention and TAU conditions, and support the link between the specific features of the intervention and the outcomes. Another limitation is that many children did not benefit from the program and for those who did, gains did not register in teacher ratings. The lack of change in teacher ratings has been seen in other studies when parent ratings have shown improvement, but it highlights the need for additional assessments of real-world function. The response rate was only 41% in the overall sample and 49% in the inattentive and combined subtypes, roughly twice the 20% during the TAU control condition. It should be noted, however, that these treatment and control response rates are similar to those of many current treatments in psychiatry, as response rates are often ~40–50% and less than twice that seen in placebo comparison groups (e.g. Findling et al., 2015; Walsh, Seidman, Sysko, & Gould, 2002). It is also consistent with the heterogeneous nature of the ADHD clinical diagnosis. Still, there is need to improve the treatment and to find different treatments for other subgroups. While the dropout rate during the intervention period was only 15%, suggesting feasibility and general acceptance of the intervention, when dropouts during TAU are added the total reaches nearly 30% urging further caution in conclusions about possible impact on the overall population of ADHD. Finally, we do not have data on the durability of clinical or cognitive improvements, or the need or value of extended or booster training.

Conclusion and future directions

This study, along with recent ERP (Smith et al., 2019) and fMRI (Rosa et al., 2019) studies of the intervention used in this study, provides promising evidence that a substantial number of children with ADHD might respond to a non-pharmacologic, multidimensional cognitive-training treatment. Additional studies are needed to: (1) confirm or repudiate this finding; (2) better characterize those children likely to respond to cognitive-training; (3) determine whether modifications in the cognitive-training program can increase the degree of benefit and the number of children who respond; (4) assess the durability of improvement and value of maintenance treatment; (5) compare effectiveness of the intervention to established pharmacotherapy, and (6) evaluate the relative importance of the different components of the intervention.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291720000288

Acknowledgements. The authors would like to thank the personnel of the Hamden Public Schools for their invaluable contribution to this project. In particular, we are grateful to Robin Riccitelli, the Coordinator for Elementary Special Education Services, and Chris Brown (her predecessor) as well as the more than 20 teachers and staff members who implemented this program in their schools.

Financial support. This work was supported by a grant from the National Institute of Health to BEW and JFL (TR01 HD070821-01)

Conflict of interest. Bruce Wexler is Chief Scientist and an equity holder in the Yale Start Up company C8 Sciences that sells the cognitive training program evaluated in this study. Jinxia Dong is also an equity holder in C8 Sciences.

References

Abikoff, H., Gittelman-Klein, R., & Klein, D. F. (1977). Validation of a classroom observation code for hyperactive children. *Journal of Consulting and Clinical Psychology*, *45*, 772–783.

Abikoff, H., Hechtman, L., Klein, R. G., Weiss, G., Fleiss, K., Etcovitch, J., … Pollack, S. (2004). Symptomatic improvement in children with ADHD treated with long-term methylphenidate and multimodal psychosocial treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*, 802–811.

ADHD-200 Consortium (2012). The ADHD-200 consortium: A model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience*, *6*, Article 62, 1–5.

Allen, A. J., Kurlan, R. M., Gilbert, D. L., Coffey, B. J., Linder, S. L., Lewis, D. W., & Spencer, T. J. (2005). Atomoxetine treatment in children and adolescents with ADHD and comorbid tic disorders. *Neurology*, *65*, 1941–1949.

Amador-Campos, J. A., Forns-Santacana, M., Guàrdia-Olmos, J., & Peró-Cebollero, M. (2006). DSM-IV attention deficit hyperactivity disorder symptoms: Agreement between informants in prevalence and factor structure at different ages. *Journal of Psychopathology and Behavioral Assessment*, *28*, 23–32.

Barkley, R. A. (1977). A review of stimulant drug research with hyperactive children. *Journal of Child Psychology and Psychiatry*, *18*, 137–165.

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

Barkley, R. A. (2003). Issues in the diagnosis of attention-deficit/hyperactivity disorder in children. *Brain and Development*, *25*, 77–83.

Barkley, R. A., Anastopoulos, A. D., Guevremont, D. C., & Fletcher, K. E. (1992). Adolescents with attention deficit hyperactivity disorder: Mother-adolescent interactions, family beliefs and conflicts, and maternal psychopathology. *Journal of Abnormal Child Psychology*, *20*, 263–288.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1990). Comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. *Journal of Consulting and Clinical Psychology*, *58*, 775–789.

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

Bloch, M. H., Landeros-Weisenberger, A., Kelmendi, B., Coric, V., Bracken, M. B., & Leckman, J. F. (2006). A systematic review: Antipsychotic augmentation with treatment refractory obsessive-compulsive disorder. *Molecular Psychiatry*, *11*, 622–632.

Bussing, R., Fernandez, M., Harwood, M., Wei, H., Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: Psychometric properties and normative ratings from a school district sample. *Assessment*, *15*, 317–328.

Chacko, A., Bedard, A. C., Marks, D. J., Feirsen, N., Uderman, J. Z., Chimiklis, A., & Ramon, M. (2014). A randomized clinical trial of Cogmed working memory training in school-age children with ADHD: A replication in a diverse sample using a control condition. *Journal of Child Psychology and Psychiatry*, *55*, 247–255.

Chacko, A., Feirsen, N., Bedard, A.-C., Marks, D., Uderman, J. Z., & Chimiklis, A. (2013). Cogmed working memory training for youth with ADHD: A closer examination of efficacy utilizing evidence-based criteria. *Journal of Clinical Child and Adolescent Psychology*, *42*, 769–783.

Cherkasova, M., Sulla, E. M., Dalena, K. L., Pondé, M. P., & Hechtman, L. (2013). Developmental course of attention deficit hyperactivity disorder and its predictors. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *22*, 47–54.

Clarke, A. R., Barry, R. J., Dupuy, F. E., Heckel, L. D., McCarthy, R., Selikowitz, M., & Johnstone, S. J. (2011). Behavioural differences between EEG-defined subgroups of children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, *122*, 1333–1341.

Coghill, D. R., Seth, S., & Mathews, K. (2014). A comprehensive assessment of memory. Delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: Advancing beyond the three-pathway models. *Psychological Medicine*, *44*, 1989–2001.

Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., & Sonuga-Barke, E. J. S. (2015). Cognitive training for attention-deficit/hyperactivity disorder: Meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 164–174.

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

Dalsgaard, S., Østergaard, S. D., Leckman, J. F., Mortensen, P. B., & Pedersen, M. G. (2015). Mortality in children, adolescents, and adults with attention deficit hyperactivity disorder: A nationwide cohort study. *The Lancet*, *385*, 2190–2196.

Dovis, S., Van der Oord, S., Wiers, R. W., & Prins, P. J. M. (2015). Improving executive functioning in childen with ADHD: Training multiple executive functions within the context of a computer game. A randomized double-blind placebo controlled trial. *PLoS ONE*, *10*(4), e0121651. doi: 10.1371/journal.pone.0121651.

Dwan, K., Tianjing, L., Altman, D. G., & Elbourne, D. (2019). CONSORT 2010 Statement: Extension to randomized crossover trials. *British Medical Journal*, *366*, l4378.

Efron, D., Jarman, J., & Barker, M. (1997). Methylphenidate versus dexamphetamine in children with attention deficit hyperactivity disorder: A double-blind, crossover trial. *Pediatrics*, *100*, 1–7.

Embry, D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical Child and Family Psychology Review*, *5*, 273–297.

Erhardt, D., & Hinshaw, S. P. (1994). Initial sociometric impressions of attention-deficit hyperactivity disorder and comparison boys: Predictions from social behaviors and from nonbehavioral variables. *Journal of Consulting and Clinical Psychology*, *62*, 833–842.

Fair, D. A., Bathula, D, Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth informheterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States*, *109*, 6769–6774.

Faraone, S. V., Sergeant, J., Gillberg, C. G., & Biederman, J. (2003). The worldwide prevalence of ADHD: Is it an American condition? *World Psychiatry*, *2*, 104–113.

Findling, R. L., Landbloom, R. L., Szegedi, A., Koppenhaver, J., Braat, S., Zhu, Q., & Mathews, M. (2015). Asenapine for the acute treatment of pediatric manic of mixed episode of bipolar 1 disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 1032–1041.

Fischer, M., Barkley, R. A., Fletcher, K. E., & Smallish, L. (1993). The stability of dimensions of behavior in ADHD and normal children over an 8-year followup. *Journal of Abnormal Child Psychology*, *21*, 315–337.

Fried, R., Hirshfeld-Baker, D., Petty, C., Batchelder, H., & Biederman, J. (2015). How informative is the CANTAB to assess executive functioning in children with ADHD? A controlled study. *Journal of Attention Disorders*, *19*, 468–475.

Gray, S. A., Chaban, P., Martinussen, R., Goldberg, R., Gotlieb, H., Kronitz, R., & Tannock, R. (2012). Effects of a computerized working memory training program on working memory, attention, and academics in adolescents with severe LD and comorbid ADHD: A randomized controlled trial. *Journal of Child Psychology and Psychiatry*, *53*, 1277–1284.

Green, C. T., Long, D. L., Green, D., Iosif, A. M., Dixon, J. F., Miller, M. R., & Schweitzer, J. B. (2012). Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics*, *9*, 639–648.

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliott, J. G. (2010). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology*, *24*, 827–836.

Huizink, A. C., Van Lier, P. A. C., & Crijnen, A. A. M. (2008). Attention deficit hyperactivity disorder symptoms mediate early-onset smoking. *European Addiction Research*, *15*, 1–9.

Hyman, S. W. (2010). The diagnosis of mental disorders: The problem of reification. *Annual Reviews of Clinical Psychology*, *6*, 155–179.

Jensen, P. (2003). Longer term effects of stimulant treatments for attention-deficit/hyperactivity disorder. *Journal of Attention Disorders*, *6*, 45–56.

Johnstone, S. J., Roodenrys, S., Blackman, R., Johnston, E., Loveday, K., Mantz, S., & Barratt, M. F. (2012). Neurocognitive training for children with and without AD/HD. *ADHD Attention Deficit and Hyperactivity Disorders*, *4*, 11–23.

Karalunas, S. L., Fair, D., Musser, E. D., Aykes, K., Iyer, S. P., & Nigg, J. T. (2014). Subtyping attention-deficit/hyperactivity disorder using temperament dimensions: Toward biologically based nosologic criteria. *JAMA psychiatry*, *71*, 015–1024.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman brief intelligence test* (2nd ed.). Bloomington, MN: Pearson.

Kavanaugh, B., Tuncer, O. F., & Wexler, B. E. (2019). Measuring and improving executive function in the classroom. *Journal of Cognitive Enhancement*, *3*(3), 271–280. doi: 10.1007/s41465-018-0095-y.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD – a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 177–186.

Martel, M. M., Von Eye, A., & Nigg, J. T. (2010). Revisiting the latent structure of ADHD: Is there a 'g' factor? *Journal of Child Psychology and Psychiatry*, *51*, 905–914.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270–291.

Miller, G. (2010). Psychiatry. Beyond DSM: Seeking a brain-based classification of mental illness. *Science*, *327*, 1437.

Morimoto, S. S., Wexler, B. E., Liu, J., Hu, W., Seirup, J., & Alexopoulos, G. S. (2014). Neuroplasticity-based computerized cognitive remediation for treatment resistant geriatric depression. *Nature Communications*, *5*, 4579, 1–7. doi:10.1038/ncomms5579.

Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit-hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, *57*, 1224–1230.

Pallanti, S., Querciol, L., Sood, E., & Hollander, E. (2002). Lithium and valproate treatment of pathological gambling: A randomized single-blind study. *Journal of Clinical Psychiatry*, *63*, 559–564.

Popper CW (1988). Disorders usually first evident in infancy, childhood, or adolescence. In (J. A. Talbot, R. E. Hales, & S. C. Yudofsky (Eds.), *Textbook of psychiatry* (pp. 649–735). Washington, DC: American Psychiatric Press.

Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review*, *33*, 1237–1252.

Riddle M and Science of Behavior Change Working Group (2015). News from the NIH: Using an experimental medicine approach to facilitate translational research. *TBM*, *5*, 486–488.

Rosa, V. O., Franco, A. R., Salum, G. A., Moreora-Maia, C. R., Wagner, F., Simioni, A., & Rhode, L. A. P. (2019). Effects of computerized cognitive training as ass-on treatment to stimulants in ADHD: A pilot fMRI study. *Brain Imaging and Behavior*, 1–12. doi:10.1007/s11682-019-00137-0.

Schachar, R. (2014). Genetics of attention deficit hyperactivity disorder (ADHD): Recent updates and future prospects. *Current Developmental Disorders Reports*, *1*, 41–49.

Shalev, L., Tsal, Y., & Mevorach, C. (2007). Computerized progressive attentional training (CPAT) program: Effective direct intervention for children with ADHD. *Child Neuropsychology*, *13*, 382–388.

Sibley, M. H., Altszuler, A. R., Morrow, A. S., & Merrill, B. M. (2014). Mapping the academic problem behaviors of adolescents with ADHD. *School Psychology Quarterly*, *29*, 422–437.

Smith, S. D., Crowley, M. J., Ferrey, A., Ramsey, K., Wexler, B. E., Leckman, J. F., & Sukhodolsky, D. G. (2019). Effects of integrated brain, body, and social (IBBS) intervention on ERP measures of attentional control in children with ADHD. *Psychiatry Research*, *278*, 248–257.

Smith, S. D., Vitulano, L. A., Katsovich, L., Li, S., Moore, C., Li, F., & Leckman, J. F. (2016). A randomized controlled trial of an integrated brain, body, and social intervention for children with ADHD. *Journal of Attention Disorders*, 1–15. doi: 10.1177/1087054716647490

Sonuga-Barke, E., Bitsakou, P., & Thompson, M. (2010). Beyond the dual pathway model: Evidence for the dissociation of timing, inhibitory, and delay-related impairments in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(4), 345–355.

Sonuga-Barke, E. J., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., & Dittmann, R. W. (2013). Nonpharmacological interventions for ADHD: Systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry*, *170*, 275–289.

Storch, E. A., Lewin, A. B., De Nadai, A. S., & Murphy, T. K. (2010). Defining treatment response and remission in obsessive-compulsive disorder: A signal detection analysis of the children's Yale-Brown Obsessive Compulsive

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press

Scale. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 708–717.

Suzuki, T., Remington, G., Mulsant, B. H., Uchida, H., Rajji, T. K., Graff-Guerrero, A., & Mamo, D. C. (2012). Defining treatment-resistant schizophrenia and response to antipsychotics: A review and recommendation. *Psychiatry research*, *197*, 1–6.

Swanson, J. M. (1992). *School-Based assessments and interventions for ADD students*. Irvine, CA: KC Publishing.

van Dongen-Boomsma, M., Vollebregt, M. A., Slaats-Willemse, D., & Buitelaar, J. K. (2013). A randomized placebo-controlled trial of electroencephalographic (EEG) neurofeedback in children with attention-deficit/hyperactivity disorder. *Journal of Clinical Psychiatry*, *74*, 821–827.

Wahkstedt, C., Thorell, L. B., & Bohlin, G. (2008). ADHD symptoms and executive function impairment: Early predictors of later behavioral problems. *Developmental Psychology*, *33*, 160–178.

Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression. *JAMA*, *287*, 1840–1847.

Wexler, B. E. (1992). Beyond the Kraepelinean dichotomy. *Biological Psychiatry*, *31*, 539–541.

Wexler, B. E., & Bell, M. D. (2005). Cognitive remediation and vocational rehabilitation for schizophrenia. *Schizophrenia Bulletin*, *31*, 931–941.

Whalen, C. K., Henker, B., & Dotemoto, S. (1980). Methylphenidate and hyperactivity: Effects on teacher behaviors. *Science*, *208*, 1280–1282.

Wilens, T. E., Biederman, J., Brown, S., Monuteaux, M., Prince, J., & Spencer, T. J. (2002). Patterns of psychopathology and dysfunction in clinically referred preschoolers. *Journal of Developmental and Behavioral Pediatrics*, *23*, S31–S36.

Wykes, T., Huddy, V., Cellard, C., McGurk, S. R., & Czobor, P. (2011). A meta-analysis of cognitive remediation for schizophrenia: Methodology and effect sizes. *The American Journal of Psychiatry*, *168*, 472–485.

https://doi.org/10.1017/S0033291720000288 Published online by Cambridge University Press