Computerized Training of Working Memory in Children With ADHD—A Randomized, Controlled Trial

TORKEL KLINGBERG, M.D., PH.D., ELISABETH FERNELL, M.D., PH.D., PERNILLE J. OLESEN, M.Sc., MATS JOHNSON, M.D., PER GUSTAFSSON, M.D., PH.D., KERSTIN DAHLSTRÖM, M.D., CHRISTOPHER G. GILLBERG, M.D., PH.D., HANS FORSSBERG, M.D., PH.D., AND HELENA WESTERBERG, L.P., PH.D.

ABSTRACT

Objective: Deficits in executive functioning, including working memory (WM) deficits, have been suggested to be important in attention-deficit/hyperactivity disorder (ADHD). During 2002 to 2003, the authors conducted a multicenter, randomized, controlled, double-blind trial to investigate the effect of improving WM by computerized, systematic practice of WM tasks. **Method:** Included in the trial were 53 children with ADHD (9 girls; 15 of 53 inattentive subtype), aged 7 to 12 years, without stimulant medication. The compliance criterion (>20 days of training) was met by 44 subjects, 42 of whom were also evaluated at follow-up 3 months later. Participants were randomly assigned to use either the treatment computer program for training WM or a comparison program. The main outcome measure was the span-board task, a visuospatial WM task that was not part of the training program. **Results:** For the span-board task, there was a significant treatment effect both post-intervention and at follow-up. In addition, there were significant effects for secondary outcome tasks measuring verbal WM, response inhibition, and complex reasoning. Parent ratings showed significant reduction in symptoms of inattention and hyperactivity/impulsivity, both post-intervention and at follow-up. **Conclusions:** This study shows that WM can be improved by training in children with ADHD. This training also improved response inhibition and reasoning and resulted in a reduction of the parent-rated inattentive symptoms of ADHD. *J. Am. Acad. Child Adolesc. Psychiatry,* 2005;44(2):177–186.

Key Words: attention-deficit/hyperactivity disorder, intervention, working memory, response inhibition.

Attention-deficit/hyperactivity disorder (ADHD) affects 3% to 5% of school-age children with serious impairments in both academic performance and social functioning. Many of these problems persist into adulthood (Biederman et al., 2000; Rasmussen and Gillberg, 2000). Deficits in executive functioning, including working memory (WM), response inhibition, and temporal processing, have been suggested to play an important role in ADHD (Barkley, 1997; Castellanos and Tannock, 2002; Rapport et al., 2000). Executive functions is a broad concept that includes, among other functions, the ability to inhibit a prepotent response, planning, reasoning, and WM. WM is the ability to retain information during a delay and then to make a response based on that internal representation. Furthermore, WM is often regarded as a more fundamental function, underlying other executive functions such as reasoning. WM deficits in ADHD have been demonstrated repeatedly (Dowson

Accepted September 14, 2004.

Drs. Klingberg, Fernell, Forssberg, and Westerberg and Ms. Olesen are with the Unit of Neuropediatrics, Department Women and Children's Health, Karolinska Institute, Stockholm Institute, Stockholm; Drs. Johnson and Gillberg are with the Department of Child and Adolescent Psychiatry, Göteborg University, Sweden; Dr. Gustafsson is with the Division of Child and Adolescent Psychiatry, Faculty of Health Sciences, Linköping University, Sweden; Dr. Dahlström is with the Department of Neuropediatrics, Huddinge University Hospital, Sweden.

This study was supported by the Swedish Research Foundation (Vetenskapsrådet), the Wallenberg Global Learning Network, and Cogmed Cognitive Medical Systems AB. The authors thank Anna-Karin Adler, Gunilla Berglund, and Sven Östlund for performing the neuropsychological testing, Jonas Beckeman and David Skoglund for their significant contribution to task design, Maria Andersson for managing the patient calls, and Johan Bring for statistical analyses.

Reprint requests to Dr. Klingberg, Unit of Neuropediatrics, Department of Women and Children's Health, Astrid Lindgren's Children's Hospital, Q2:07, Karolinska Institute, 171 76 Stockholm, Sweden; e-mail: torkel.klingberg@kbh.ki.se.

et al., 2004; Karatekin and Asarnow, 1998; Kempton et al., 1999; Kuntsi et al., 2001; Mariani and Barkley, 1997; Westerberg et al., 2004, but see also Karatekin, 2004).

This study investigated whether systematic training of WM tasks during a 5-week period would improve WM, improve other executive functions, and reduce the ADHD symptoms. Several studies have evaluated the effect of practice with various types of cognitive tasks in subjects with stroke (Sohlberg et al., 2000), in elderly subjects (Ball et al., 2002), and after traumatic brain injury (Salazar et al., 2000). The method evaluated in this study differs from that of previous ones in that it focuses entirely on training WM tasks. Moreover, the training is computerized, which makes it possible to automatically and continuously adapt the difficulty level to the performance of the child to optimize the training effect.

The effect of WM training on brain activity was recently evaluated with functional magnetic resonance imaging (Olesen et al., 2004). In that study, young, healthy adult subjects were scanned while performing a WM task and a control task before and after WM training. Training improved the WM performance of the subjects and resulted in increased brain activity in the dorsolateral prefrontal and parietal association cortices, indicating plasticity of the neural systems underlying WM. These cortical areas partly overlap with the prefrontal regions implicated in ADHD pathology (Castellanos et al., 1996, 2002; Filipek et al., 1997), and this provides a neuroanatomical rationale for undertaking WM training in children with ADHD. A previous preliminary study indicated that training of WM tasks can enhance executive functioning including WM, response inhibition, and reasoning in children with ADHD (Klingberg et al., 2002b). A major shortcoming of that study was the low number of subjects ($n = 7$ in both the treatment and the comparison groups). Moreover, ratings of ADHD symptoms were not performed, only one clinical center was involved, and there was no follow-up measurement of both groups to estimate the extent to which training effects lasted. The current study was therefore conducted at four clinical sites evaluating the effects of practice of WM tasks in a randomized, controlled, double-blind design. Executive functions were measured and ADHD symptoms were rated before, immediately after, and 3 months after intervention.

In this study, we compared two similar versions of the same training program. In the treatment program, the children practiced WM tasks in which the difficulty level was adjusted to closely match the WM capacity of the child. This procedure was hypothesized to optimize the training effect. In the comparison condition, the same tasks were used, but the WM load (i.e., number of items to be remembered) was low, thus resulting in easy tasks that were expected to result in only small training effects. By having two similar versions, we intended to control as much as possible for nonspecific effects of the training procedure, such as expectancy, passage of time, and maturation and specifically estimate the effect of improvement of WM. To evaluate the effect of training, we used tasks that were not part of the training program.

METHOD

Subjects

Referral sources included pediatricians, child psychiatrists, and special teachers in schools. We included only nonmedicated children because they were thought to have more room for clinical improvement of ADHD symptoms than children on medication and therefore give a better chance of detecting significant treatment effects. Furthermore, in Sweden, only a minority of children with ADHD receives medication.

Inclusion criteria were (1) diagnosis of ADHD of either combined or predominantly inattentive subtype, (2) age between 7 and 12 years at inclusion, and (3) access to a personal computer with an Internet connection at home or in school. Exclusion criteria were (1) being treated with stimulants, atomoxetine, neuroleptic, or any other psychoactive drugs; (2) fulfilling criteria for diagnosis of clinically significant oppositional defiant disorder, autistic syndrome, Asperger's syndrome or depression; (3) history of seizures during the past 2 years; (4) IQ <80 (based on an IQ test or the physician's clinical impression and school history); (5) motor or perceptual handicap that would prevent using the computer program; (6) educational level and socioeconomic situation that made it unlikely that the family would be able to follow the treatment procedure and study requirements (the educational level of the parents was not specified in terms of academic degree); and (7) medical illness requiring immediate treatment.

Of 56 patients attending the screening visit, 53 were included in the study and randomized to the treatment or comparison programs (Table 1, Fig. 1). Diagnostic assessment, including subtyping, had in most cases been made before the screening visit but was confirmed by the physician and based on global clinical impression taking *DSM-IV* rating scales (American Psychiatric Association, 1994) from parents and teachers into account. None of the children fulfilled criteria for conduct disorder or bipolar disorder. Two of the children had previously taken stimulant medication but stopped more than 1 year before the study. One child discontinued stimulant medication 1 week before the first measurements to be able to participate in the study. All other children had never been on medication for ADHD.

The study was approved by the regional ethics committee at Karolinska Hospital and by the local ethics committees at the four

TABLE 1
Subject Characteristics[a]

	Comparison	Treatment	Total
Boys	22/20	22/16	44/36
Girls	4/4	5/4	9/8
ADHD combined	16/15	22/15	38/30
ADHD inattentive	10/9	5/5	15/14
Age, yr, mean (SD)	9.8 (1.3)/9.7 (1.3)	9.9 (1.3)/9.8 (1.4)	9.8 (1.3)/9.8 (1.3)

Note: ADHD = attention-deficit/hyperactivity disorder.
[a] Data given for all randomized subjects (n = 53)/subjects that complied (n = 44). See Fig. 1.

participating sites. Written informed consent was obtained from all participating families.

Outcome Measures

Certified psychologists performed neuropsychological assessments at the four clinical sites using four tasks to evaluate executive performance. (1) The span-board task from the WAIS-RNI testing battery (Wechsler, 1981) was used to measure visuospatial WM. The mean performance from trials with forward and backward repeating of the memoranda was used in the analysis to provide a more reliable measure. (2) Digit-span from the WISC-III testing battery was used to measure verbal WM. (3) The Stroop interference task was used to measure response inhibition (Lezak, 1995). In this task words describing colors are printed with ink in a color that was incongruent with the word, i.e., "green" printed in yellow ink. The subjects were asked to name the color of the ink for each word. (4) Raven's Colored Progressive Matrices (Raven, 1995) was used to measure nonverbal reasoning ability. Motor activity was measured by an infrared camera that records the number of head movements during 15 minutes of performance of a detection task on a computer (Teicher et al., 1996).

The 18 *DSM-IV* items were used as a rating scale for ADHD symptoms (American Psychiatric Association, 1994). We also used the Conners Rating Scale for parents and teachers (revised, short version) (Conners, 2001). Symptoms were scored by assigning a severity estimate for each symptom on a 4-point scale, from 0 (not at all) to 3 (very much). Because there are no updated Swedish norms for these scales, raw values were used and reported.

Before the study was conducted, the span-board task was defined to be the main outcome measure because it provides a nontrained measure of visuospatial WM. The stimuli, presentation, and response mode for this task differ from the WM tasks that were part of the training program (see Discussion).

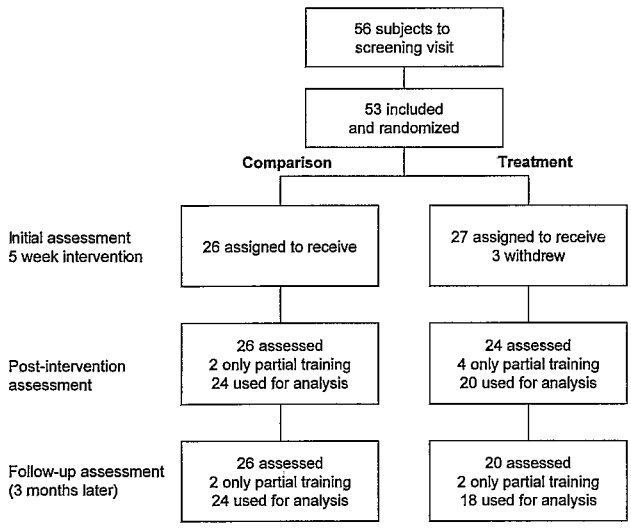

Fig. 1 Flow of participants through the trial.

Intervention

The treatment consisted of performing WM tasks implemented in a computer program developed for this study (RoboMemo®, Cogmed Cognitive Medical Systems AB, Stockholm, Sweden). The program was provided on a CD and used by the child on a personal computer either at home or in school. The program included visuospatial WM tasks (remembering the position of objects in a 4 × 4 grid as well as verbal tasks (remembering phonemes, letters, or digits) (see Olesen et al., 2004) for further description of the visuospatial tasks). Responses were made by clicking on displays with the computer mouse. The children performed 90 WM trials on each day of training. Total time depended on the level and time between trials. Medium total training time (excluding breaks) was about 40 minutes. The difficulty level was automatically adjusted, on a trial-by-trial basis, to match the WM span of the child on each task. Responses to each trial were logged to a file on the computer, and every 1 to 2 days, an adult used a report program to upload the log file via the Internet to a server so that compliance could be verified.

The comparison condition was identical to the treatment except that the difficulty of the 90 WM trials remained on the initial low level (two to three items) instead of being increased to match the WM span of the child. As in the treatment program, responses were logged to a file that was uploaded to a server to verify compliance.

Procedures

Children were recruited via pediatricians, child psychiatrists, and special teachers. Initial information was given verbally. Families expressing interest in participating in the study were sent the detailed written information about the study, together with *DSM-IV* questions about ADHD and oppositional defiant disorder, and Conners rating scales, to be answered before the screening visit. The written information stated that there would be two interventions and that one of them probably was more effective than the other. The written information also specified the inclusion criteria regarding diagnosis, medication, age, and the need for a computer with Internet access. At

the screening visit, the physician met both the child and one parent and checked all inclusion and exclusion criteria. The included children were scheduled for baseline evaluation (= time point 1 [T1]) at which a psychologist administered the tests and provided the CD with the computer program. They were shown how to install the program and asked to complete at least 25 days of training before the next visit.

A randomized, blinded list of numbers associated with the CDs containing the treatment or comparison program was sent to each clinical center. Randomization was done with blocks of four. The CDs were distributed by the testing psychologists to the children in the order that they entered the study at each site. Thus, the physician, psychologist, parent, and child were all blinded to child group status until after the follow-up assessment. Once every week, a certified psychologist (not affiliated with any of the testing sites) called the parents and asked about technical difficulties and gave feedback about how many days of training that had been uploaded to the server. This was intended to increase compliance. The postintervention visit (T2) took place 5 to 6 weeks after the baseline visit, and the follow-up assessment (T3) was done 3 months after T2. On both occasions, the psychologist again administered the neuropsychological tests and distributed rating scales that were completed at home by the parent and teacher and mailed to the study coordinator. The study was conducted during 2002, with follow-up during 2003.

Statistical Analysis

Hypotheses were tested by comparing outcome score at later times (T2 or T3) for the two groups using a general linear model, controlling for age, number of days of program use, and baseline score (T1). This analysis is equivalent to a between-group analysis of covariance with baseline as a covariate. We also controlled for the location where the intervention took place (home or school) when analyzing ratings from teacher and parent. Analyses showed no effect of gender or clinical site for any of the variables, and these two covariates were therefore not included in the further analyses. One-tailed tests were used because all hypotheses were based on the preliminary study (Klingberg et al., 2002b), and only the superiority of treatment over comparison was of interest. Analyses were done using SAS v8.02 (SAS Institute Inc., Cary, NC).

RESULTS

The flow of participants is shown in Figure 1. Of the 53 subjects, three withdrew: two because of computer problems and one because of social problems not related to the study. Fifty children attended the postintervention visit. The criterion for sufficient compliance was defined before the study to be 20 or more days of program use. Forty-four of the 50 subjects met these criteria. Mean number of days was 26.6 (SD = 2.6) in the comparison group and 25.2 (SD = 2.2) in the treatment group.

In the analysis of rating scales, missing values were not imputed in the regression analysis. Hence, only patients with complete data were included in the model. The main reason for missing values was that parents or teachers failed to mail the rating scales to the study coordinator. The regression analyses of rating scale data were thus based on a smaller number of observations: 36 from parents, 34 from teachers at T2, and 37 from parents and 35 from teachers at T3. To check for any selection bias, we compared patients with complete rating data with those with missing data and found no significant differences in baseline values from the rating scales. Subjects with missing data were also evenly distributed between the two groups (Tables 1 and 2).

Executive Tasks

Raw data (not corrected for the covariates) are shown in Table 2. Mean values corrected for covariates (including baseline measurement) are shown in Figure 2. The results of the statistical analysis are shown in Table 3. For the main outcome measure, the span-board task, there was a significantly greater improvement from baseline to postintervention measurement in the treatment group compared with the comparison group (Fig. 2A, Table 3). At follow-up, the treatment effect was still significant. The effect size (Cohen's delta [Cohen, 1988]) was calculated from corrected values (Fig. 2, Tab. 3) and was 0.93 post-intervention and 0.92 (99%) at follow-up.

Other executive tasks were considered secondary outcome measures. There was a significant treatment effect for all executive tasks (Table 3). The faster performance on the Stroop task was not due to a prioritizing of speed while sacrificing accuracy because the treatment group was not only faster but also more accurate (Table 2). Effect sizes were 0.59, 0.34, and 0.45 for the digit-span, Stroop time, and Raven's task, respectively.

At follow-up, the performance in the treatment group was about as high as, or higher than, at postintervention (Table 2, Fig. 2B–D). The effect sizes were 0.57, 0.25, and 0.30 for the digit-span, Stroop time, and Raven's task, respectively. This corresponds to 97%, 73%, and 67%, respectively, of the postintervention effect. The ability to detect the remaining significant effect for the Raven's task and for accuracy in the Stroop task was limited by the ceiling effects, which were more pronounced for the treatment group. Postintervention, eight subjects (40%) in the treatment group had a score of 34 to 36 on the Raven's task, for which 36 is the maximal score. In the comparison group, the corresponding number was 1 (4%). There was neither significant baseline-score-by-treatment interaction nor any significant ADHD-subtype-by-treatment interaction

TABLE 2
Mean Values for the Neuropsychological Assessment and Rating Scales

	Baseline	Post-intervention		Follow-up	
	Mean (SD) [*n*]	Mean (SD) [*n*]	SC	Mean (SD) [*n*]	SC
Span-board (items) (visual WM)					
c	4.13 (0.75) [24]	4.28 (0.81) [24]	0.20	4.35 (0.86) [24]	0.29
t	4.40 (0.95) [20]	5.22 (1.01) [20]	0.86	5.31 (1.05) [18]	0.96
Digit-span (items) (verbal WM)					
c	3.69 (0.80) [24]	3.73 (0.73) [24]	0.05	3.64 (0.65) [23]	−0.06
t	3.62 (0.63) [20]	4.08 (0.89) [20]	0.73	4.01 (1.03) [18]	0.62
Stroop accuracy (max 60) (response inhibition)					
c	54.6 (3.69) [23]	55.1 (4.77) [24]	0.13	57.6[a] (2.38) [24]	0.81[a]
t	55.8 (4.45) [20]	58.3[a] (1.84) [20]	0.58[a]	58.1[a] (2.60) [18]	0.54[a]
Stroop time (s) (response inhibition)					
c	121.5 (28.6) [23]	117.9 (35.1) [24]	−0.13	109.5 (32.6) [24]	−0.41
t	123.0 (56.3) [20]	101.8 (31.6) [24]	−0.38	94.2 (27.9) [18]	−0.51
Raven accuracy (max 36) (reasoning)					
c	25.3 (4.0) [24]	26.5 (4.1) [24]	0.30	27.7 (4.4) [24]	0.59
t	28.8 (4.6) [20]	31.0[a] (4.6) [20]	0.48[a]	31.4[a] (5.0) [18]	0.56[a]
Head movements					
c	2,156 (1,452) [23]	2,404 (1,607) [23]	0.17	2,362 (1,709) [23]	0.14
t	1,932 (1,105) [19]	2,304 (1,738) [20]	0.34	1,770 (1,201) [16]	−0.15
ADHD parent inatt.					
c	16.6 (5.2) [23]	15.5 (7.2) [19]	0.21	14.5 (5.6) [20]	0.40
t	18.7 (5.1) [20]	13.0 (6.5) [17]	1.12	13.6 (7.9) [17]	1.01
ADHD Parent H/I					
c	13.9 (7.7) [23]	12.5 (7.0) [19]	0.18	13.0 (7.4) [20]	0.12
t	13.0 (8.8) [20]	9.8 (7.8) [17]	0.36	8.6 (8.0) [17]	0.50
ADHD teacher inatt.					
c	13.1 (7.6) [23]	12.7 (7.1) [20]	0.05	12.9 (6.5) [20]	0.03
t	16.6 (5.7) [20]	14.6 (5.8) [14]	0.35	15.6 (6.7) [15]	0.17
ADHD teacher H/I					
c	11.1 (6.4) [23]	10.3 (6.9) [20]	0.13	10.4 (7.1) [20]	0.10
t	16.2 (9.5) [20]	14.1 (8.7) [14]	0.22	15.9 (8.0) [15]	0.03

Note: WM = working memory; c = comparison; t = treatment; SC = standardized change; SC_{T2}= ($mean_{T2}$ − $mean_{T1}$)/SD_{T1}; inatt = symptoms of inattention (high score = more inattentive); H/I = hyperactivity/impulsivity score; ADHD = attention-deficit/hyperactivity disorder.
[a] Ceiling effects limit the interpretability of these scores. All values are raw scores, uncorrected for any covariate.

for any variable (neither for the executive tasks or the rating scales). All significant differences remained significant also when ADHD subtype was included as a covariate in the analyses.

Ratings of ADHD Symptoms

In the parent ratings of ADHD symptoms (according to *DSM-IV*), there was a significant reduction in the inattention and hyperactivity/impulsivity scores (Table 3). The rating scores from the teachers were not significant (Table 2). Effect sizes (from corrected values) for the significant comparisons were 1.21 for parent-rated attention at T2, and 0.67 at T3. For parent-rated hyperactivity + impulsivity, it was 0.42 both at T2 and T3. It should be noted that these effect sizes are based on differences relative to a comparison group that was not passive but received a form of low-dose treatment.

From Conners rating scales, the Parent Rating Scale showed a significant decrease for parent ratings of oppositional symptoms at follow-up (n = 35, p = .02) of hyperactivity symptoms, post-intervention (n = 36, p = .03), and at follow-up (n = 35, p = .002) and of the ADHD index at follow-up (n = 35, p = .02). Other comparisons were not significant.

Noncompliers

To check for possible effects of the subjects who were excluded because of noncompliance (<20 days of practice),

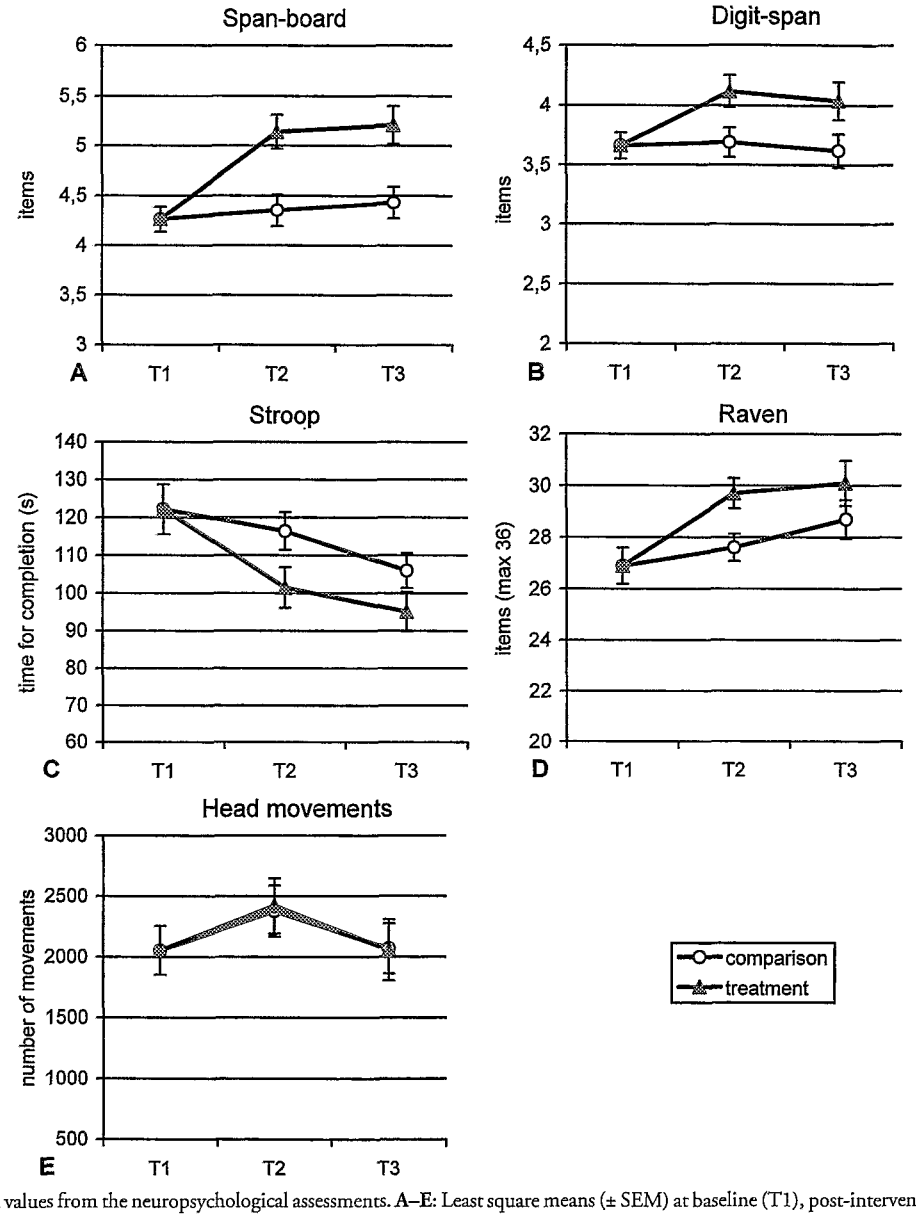

Fig. 2 Corrected mean values from the neuropsychological assessments. **A–E:** Least square means (± SEM) at baseline (T1), post-intervention (T2), and follow-up (T3) corrected for differences in baseline score, with the baseline score set to the common average for both groups.

we also performed an analysis in which we included all subjects for whom we had outcome measures at T2 ($n = 50$). All tests that were significant in the first analysis (span-board, digit-span, Stroop time, Raven's task, parent ratings of attention and hyperactivity) were also significant when the additional six subjects were included ($p < .05$).

Adverse Events

At the time of testing after intervention and at follow-up, parents and children were asked about any adverse events that could be related to the intervention. There were none.

DISCUSSION

In this study, the treatment group that undertook high-intensity training of WM improved significantly more than the comparison group on the main outcome measure: the span-board task, which was a nonpracticed measure of visuospatial WM. This effect also remained at follow-up. In addition, there were treatment effects

TABLE 3
General Linear Model Analysis

	Post-intervention				Follow-up			
	n	R^2	β	*p*	*n*	R^2	β	*p*
Cognitive tasks								
Span-board task	44	0.49	0.79	.001	42	0.49	0.78	.002
Digit-span task	44	0.51	0.43	.01	41	0.45	0.42	.03
Raven's task	44	0.77	2.1	.01	42	0.57	1.4	.12
Stroop (accuracy)	43	0.19	3.5	.004	42	0.23	0.1	.44
Stroop (time)	43	0.56	−14.9	.025	42	0.49	−11	.07
Motor activity (movements)	42	0.70	41	.90	37	0.62	−9	.97
ADHD symptoms ratings								
Parent inattention	36	0.53	−6.3	.002	37	0.50	−3.5	.04
Parent hyp/imp	36	0.69	−3.4	.03	37	0.73	−3.4	.03
Teacher inattention	33	0.61	−0.6	.37	34	0.40	1.5	.46
Teacher hyp/Imp	33	0.71	1.0	.58	34	0.68	1.3	.56

Note: R^2 = proportion of variance explained by the model; β = treatment group coefficient showing the difference in means (treatment comparison) controlling for covariates; *p* = significance level (one-tailed); ADHD = attention-deficit/hyperactivity disorder; hyp/imp = hyperactivity/impulsivity score.

for response inhibition (Stroop task), verbal WM (digit-span), complex reasoning (Raven's task), and for parent ratings of ADHD symptoms.

The span-board task differs from the trained visuospatial WM tasks with respect to the stimuli that are used (blocks on a board versus circles lighting up and disappearing on a screen), stimulus configuration (10 irregularly positioned blocks versus 16 boxes in a regular 4 × 4 grid) as well as response mode (pointing versus using the computer mouse), and the testing situation (interacting with a person versus a computer). The improvement on the span-board task is therefore evidence that the training effect generalized to a nontrained visuospatial WM task. The treatment effect (β in Table 3) corresponds to a 19% improvement, and the effect size was 0.93. The effect of stimulant medication, by comparison, has in different studies improved visuospatial WM with effect sizes of approximately 0.5 (Barnett et al., 2001), 0.4 to 1.2 (Bedard et al., 2004), and 0.4 to 0.7 (Kempton et al., 1999). The effect of training on visuospatial WM was thus fully comparable with that of medication. Comparisons with previous studies of WM in children with and without ADHD suggests that the treatment-related improvement brought the children's WM performance as close as 0 to 0.3 SD below that of the rest of the population (Barnett et al., 2001; Kempton et al., 1999; Mariani and Barkley, 1997; Westerberg et al., 2004). Although tasks such as the span-board task are laboratory tasks and it can be difficult to see the usefulness in everyday life of an increase in WM span, it is known that WM capacity underlies a wide range of cognitive abilities, including reasoning and control of attention (Engle et al., 1999) and the ability to resist distraction from irrelevant stimuli (de Fockert et al., 2001). It has also been suggested that deficits in nonverbal WM in children with ADHD contributes to the inability to hold events in mind and to defective planning (Barkley, 1997).

The three other executive tasks (digit-span, Stroop task, and Raven's task) were secondary outcome measures, and the outcome of the statistical tests for these tasks should therefore be interpreted cautiously. However, group differences for Raven's task and the Stroop task were also found in the preliminary study of children with ADHD (Klingberg et al., 2002b) as well as in a study of WM training in adults (Olesen et al., 2004). Together, these results indicate that the effect of WM training also transfers to nontrained executive tasks other than WM tasks. The training program did not include any problem-solving task or any response inhibition task even remotely similar to the Stroop task or Raven's task. The Stroop task measures the ability to inhibit a prepotent response. A review has also shown this task to be a sensitive task for measuring cognitive impairments in ADHD (Barkley et al., 1992). Also in regard to this task, comparison with previous studies suggests that the training brought the children's performance to less than 0.3 SD from the population mean

(Nigg et al., 2002; Willcutt et al., 2001). The improvement in reasoning ability after training of WM is consistent with the psychological literature, suggesting that WM is necessary for reasoning ability (Engle et al., 1999) and that visuospatial WM correlates highly with performance on Raven's task (Fry and Hale, 1996). Neuroanatomically, this association could be explained by overlapping parts of cortex in the prefrontal and parietal lobe used for WM and reasoning (Gray et al., 2003). Visuospatial WM and response inhibition also have neuroanatomical commonalities. The same areas in the superior part of the prefrontal cortex and in the parietal cortex underlie development of both visuospatial WM capacity (Klingberg et al., 2002a) and performance of the Stroop task (Adleman et al., 2002). Overlapping neural systems could thus explain how training of visuospatial WM could generalize to reasoning and response inhibition. This interpretation is also supported by the finding that training of WM tasks increase brain activity in multimodal areas of the prefrontal and parietal cortices (Olesen et al., 2004).

The WM training had a very strong effect on parent-rated attention but not on teacher ratings. We have no clear explanation for this discrepancy at this time. We have no data on how much time that the teachers spent with the children each day or each week. It has been shown that interrater correlations are often low (Swanson et al., 2001), and a twofold difference in the effect sizes from parents and teachers is also seen in other studies (Wolraich et al., 2001). Cultural differences in teacher ratings have also been reported (Magnusson et al., 1999). This, combined with a modest number of subjects and some missing data, could result in some discrepancies just from random variability. However, many of the symptoms of inattention are closely related to executive functions and we see the significant ratings from parents as more reliable because they are consistent with the changes measured in the objective tests of executive functioning.

Although ratings from parents also showed significant effects for symptoms related to hyperactivity, this was not confirmed by any significant decrease in the number of measured head movements (Table 3, Fig. 2E). The reason for this is unclear. Additional objective measures of hyperactivity could have been useful.

Guidelines have been proposed, according to which an effect size (Cohen's delta) of 0.2 represents a weak, 0.5 a moderate, and 0.8 a strong clinical effect (Cohen, 1988; Swanson et al., 2001). By that standard, the training resulted in a strong clinical effect for the main outcome measure (span-board) as well as for parent ratings of inattentive symptoms.

Limitations

More studies will be needed to confirm the training effects in other populations and to answer additional questions about the mechanisms underlying training-induced improvements of WM. Although oppositional defiant disorder symptoms were present in many children, we excluded children fulfilling the diagnostic criteria for oppositional defiant disorder, and this limits the generalizability of the findings. A standardized psychiatric interview, as it is defined in the United States, was not performed. However, the interviews followed a prespecified protocol and all clinicians had many years of experience in diagnosing ADHD. One limitation of this study was the modest number of subjects. Additional follow-up measurements would also have been useful. In the current study, the follow-up time was limited by ethical considerations about the comparison group, who needed to be without intervention for the entire follow-up period. We have not evaluated the effect of combining medication with training, which will need to be done in the future. However, the previous preliminary study (Klingberg et al., 2002b) included children on medication. These children also improved, although the relative training effect on and off medication was not statistically evaluated because of the small sample size. Furthermore, in studies of motor training, amphetamine enhances training-induced plasticity (Butefisch et al., 2002).

Although the training effect remained relatively stable for several months, we expect that it will eventually be necessary with a shorter period of retraining to maintain the effect. However, if WM and executive functions improve by practice, as this study indicates, then we would expect some degree of practice effect also from everyday activities with very high WM loads, such as mathematics and other demanding academic activities. It is theoretically possible that improvement of WM and executive functioning by an intensive training program would enable the children to perform better and hence to participate more in such WM demanding activities. This would lead to more WM practice in everyday life, and the children would enter a positive

feedback loop that would reduce the need for retraining with a training program.

Clinical Implications

This study shows that WM can be improved by training. In addition, we saw effects on reasoning, response inhibition, and a decrease in parent-rated symptoms of ADHD. The subjects that would be expected to benefit from training of WM are presumably those individuals for whom executive deficits and inattention problems constitute a bottleneck for everyday functioning or academic performance. These individuals could be found both in the inattentive and combined subgroup of ADHD (Chhabildas et al., 2001). It is also possible that training of WM will be useful in other conditions in which WM deficits are prominent, such as after traumatic brain injury and stroke affecting the frontal lobe.

Disclosure: Drs. Forssberg and Klingberg and Ms. Westerberg own stock in Cogmed. Ms. Olesen had a consultancy agreement with Cogmed. The other authors have no financial relationships to disclose.

REFERENCES

Adleman NE, Menon V, Blasey CM et al. (2002), A developmental fMRI study of the Stroop Color-Word Task. *Neuroimage* 16:61–75

American Psychiatric Association (1994), *Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).* Washington, DC: American Psychiatric Association

Ball K, Berch DB, Helmers KF et al. (2002), Effects of cognitive training interventions with older adults: a randomized controlled trial. *JAMA* 288:2271–2281

Barkley RA (1997), Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull* 121: 65–94

Barkley RA, Grodzinsky G, DuPaul GJ (1992), Frontal lobe functions in attention deficit disorder with and without hyperactivity: a review and research report. *J Abnorm Psychol* 20:163–188

Barnett R, Maruff P, Vance A et al. (2001), Abnormal executive function in attention deficit hyperactivity disorder: the effect of stimulant medication and age on spatial working memory. *Psychol Med* 31:1107–1115

Bedard AC, Martinussen R, Ickowicz A, Tannock R (2004), Methylphenidate improves visual-spatial memory in children with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 43: 260–268

Biederman J, Mick E, Faraone SV (2000), Age-dependent decline of symptoms of attention deficit hyperactivity disorder: impact of remission definition and symptom type. *Am J Psychiatry* 157:816–818

Butefisch CM, Davis BC, Sawaki L et al. (2002), Modulation of use-dependent plasticity by d-amphetamine. *Ann Neurol* 51:59–68

Castellanos FX, Giedd JN, Marsh WL et al. (1996), Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch Gen Psychiatry* 53:607–616

Castellanos FX, Lee PP, Sharp W et al. (2002), Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *JAMA* 288:1740–1748

Castellanos FX, Tannock R (2002), Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nat Rev Neurosci* 3:617–628

Chhabildas N, Pennington BF, Willcutt EG (2001), A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *J Abnorm Child Psychol* 29:529–540

Cohen J (1988), *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Erlbaum

Conners CK (2001), *Conners Rating Scales-Revised.* North Tonawanda, NY: Multihealth Systems

de Fockert JW, Rees G, Frith CD, Lavie N (2001), The role of working memory in visual selective attention. *Science* 291:1803–1806

Dowson JH, McLean A, Bazanis E et al. (2004), Impaired spatial working memory in adults with attention-deficit/hyperactivity disorder: comparisons with performance in adults with borderline personality disorder and in control subjects. *Acta Psychiatr Scand* 110:45–54

Engle RW, Kane JM, Tuholski SW (1999), Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In: *Models of Working Memory,* Myake A, Shah P, eds. Cambridge: Cambridge University Press, pp 102–134

Filipek PA, Semrud-Clikeman M, Steingard RJ, Renshaw PF, Kennedy DN, Biederman J (1997), Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology* 48:589–601

Fry AF, Hale S (1996), Processing speed, working memory, and fluid intelligence. *Psychol Sci* 7:237–241

Gray JR, Chabris CF, Braver TS (2003), Neural mechanisms of general fluid intelligence. *Nat Neurosci* 6:316–322

Karatekin C (2004), A test of the integrity of the components of Baddeley's model of working memory in attention-deficit/hyperactivity disorder (ADHD). *J Child Psychol Psychiatry* 45:912–926

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

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

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

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

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

Lezak MD (1995), *Neuropsychological Assessment.* New York: Oxford University Press

Magnusson P, Smari J, Gretarsdottir H, Prandardottir H (1999), Attention-deficit/hyperactivity symptoms in Icelandic schoolchildren: assessment with the Attention Deficit/Hyperactivity Rating Scale-IV. *Scand J Psychol* 40:301–306

Mariani MA, Barkley RA (1997), Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. *Dev Psychol* 13:111–129

Nigg JT, Blaskey LG, Huang-Pollock CL, Rappley MD (2002), Neuropsychological executive functions and DSM-IV ADHD subtypes. *J Am Acad Child Adolesc Psychiatry* 41:59–66

Olesen P, Westerberg H, Klingberg T (2004), Increased prefrontal and parietal brain activity after training of working memory. *Nat Neurosci* 7: 75–79

Rapport MD, Chung KM, Shore G, Denney CB, Isaacs P (2000), Upgrading the science and technology of assessment and diagnosis: laboratory and clinic-based assessment of children with ADHD. *J Clin Child Psych* 29:555–568

Rasmussen P, Gillberg C (2000), Natural outcome of ADHD with developmental coordination disorder at age 22 years: a controlled, longitudinal, community-based study. *J Am Acad Child Adolesc Psychiatry* 39:1424–1431

Raven JC (1995), *Coloured Progressive Matrices.* Oxford: Oxford Psychological Press

Salazar AM, Warden DL, Schwab K et al. (2000), Cognitive rehabilitation for traumatic brain injury: a randomized trial. Defense and Veterans Head Injury Program (DVHIP) Study Group. *JAMA* 283:3075–3081

Sohlberg MM, McLaughlin KA, Pavese A, Heidrich A, Posner MI (2000), Evaluation of attention process training and brain injury education in persons with acquired brain injury. *J Clin Exp Neuropsych* 22:656–676

Swanson JM, Kraemer HC, Hinshaw SP et al. (2001), Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *J Am Acad Child Adolesc Psychiatry* 40:168–179

Teicher MH, Ito Y, Glod CA, Barber NI (1996), Objective measurement of hyperactivity and attentional problems in ADHD. *J Am Acad Child Adolesc Psychiatry* 35:334–342

Wechsler D (1981), *WAIS-R Manual.* New York: The Psychological Corporation

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

Willcutt EG, Pennington BF, Boada R et al. (2001), A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *J Abnorm Psychol* 110:157–172

Wolraich ML, Greenhill LL, Pelham W et al. (2001), Randomized, controlled trial of oros methylphenidate once a day in children with attention-deficit/hyperactivity disorder. *Pediatrics* 108:883–892

Once-Daily Atomoxetine Treatment for Children With Attention-Deficit/Hyperactivity Disorder, Including an Assessment of Evening and Morning Behavior: A Double-Blind, Placebo-Controlled Trial. Douglas K. Kelsey, MD, PhD, Calvin R. Sumner, MD, Charles D. Casat, MD, Daniel L. Coury, MD, Humberto Quintana, MD, Keith E. Saylor, PhD, Virginia K. Sutton, PhD, Jill Gonzales, BS, Sandra K. Malcolm, BS, Kory J. Schuh, PhD, Albert J. Allen, MD, PhD

Objectives: Atomoxetine seems to be as effective for treating attention-deficit/hyperactivity disorder (ADHD) when the daily dose is administered once in the morning as when the dose is divided and administered in the morning and evening. In the present study, the efficacy of atomoxetine administered once daily among children with ADHD was assessed throughout the day, including the evening and early morning. Another goal was to determine how early in treatment it was possible to discern a specific effect of the drug on ADHD symptoms. *Methods:* This study was a randomized, multicenter, double-blind, placebo-controlled trial conducted at 12 outpatient sites in the United States. A total of 197 children, 6 to 12 years of age, who had been diagnosed as having ADHD, on the basis of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.) criteria, were randomized to receive 8 weeks of treatment with atomoxetine or placebo, dosed once daily in the mornings. ADHD symptoms were assessed with parent and investigator rating scales. The primary outcome measure was the Attention-Deficit/Hyperactivity Disorder Rating Scale-IV-Parent Version: Investigator-Administered and Scored total score. Daily parent assessments of children's home behaviors in the evening and early morning were recorded with an electronic data entry system. This instrument measures 11 specific morning or evening activities, including getting up and out of bed, doing or completing homework, and sitting through dinner. *Results:* Seventy-one percent of the children enrolled were male, 69% met criteria for the combined subtype (both inattentive and hyperactive/impulsive symptoms), and the most common psychiatric comorbidity was oppositional defiant disorder (35%). Once-daily atomoxetine (final mean daily dose of 1.3 mg/kg) was significantly more effective than placebo in treating core symptoms of ADHD. Mean reductions in the Attention-Deficit/Hyperactivity Disorder Rating Scale-IV-Parent Version: Investigator-Administered and Scored total score were significantly greater for patients randomized to atomoxetine, beginning at the first visit after the initiation of treatment and continuing at all subsequent visits. Both inattentive and hyperactive/impulsive symptom clusters were significantly reduced with atomoxetine, compared with placebo. With continued treatment and dose titrations, core symptoms of ADHD continued to decrease throughout the 8-week study. Mean reductions in the daily parent assessment total scores for patients randomized to atomoxetine were superior during the first week, beginning with the first day of dosing, and were also superior at endpoint. Efficacy outcomes for the evening hours for atomoxetine-treated patients were superior to those for placebo-treated patients, as assessed with 2 different assessment scales. Decreases in the daily parent assessment morning subscores at endpoint showed a significant reduction in symptoms that lasted into the mornings. Rates of discontinuations attributable to adverse events were <5% for both groups. Adverse events reported significantly more frequently with atomoxetine were decreased appetite, somnolence, and fatigue. *Conclusions:* Among children 6 to 12 of age who had been diagnosed as having ADHD, once-daily administration of atomoxetine in the morning provided safe, rapid, continuous, symptom relief that lasted not only into the evening hours but also into the morning hours. Atomoxetine treatment was safe and well tolerated. **Pediatrics** 2004;114:e1–e8.