Training of Working Memory in Children With ADHD

Torkel Klingberg, Hans Forssberg, and Helena Westerberg

Department of Neuropediatrics, Karolinska Institute, Stockholm, Sweden

ABSTRACT

Working memory (WM) capacity is the ability to retain and manipulate information during a short period of time. This ability underlies complex reasoning and has generally been regarded as a fixed trait of the individual. Children with attention deficit hyperactivity disorder (ADHD) represent one group of subjects with a WM deficit, attributed to an impairment of the frontal lobe. In the present study, we used a new training paradigm with intensive and adaptive training of WM tasks and evaluated the effect of training with a double blind, placebo controlled design. Training significantly enhanced performance on the trained WM tasks. More importantly, the training significantly improved performance on a nontrained visuo-spatial WM task and on Raven's Progressive Matrices, which is a nonverbal complex reasoning task. In addition, motor activity – as measured by the number of head movements during a computerized test – was significantly reduced in the treatment group. A second experiment showed that similar training-induced improvements on cognitive tasks are also possible in young adults without ADHD. These results demonstrate that performance on WM tasks can be significantly improved by training, and that the training effect also generalizes to nontrained tasks requiring WM. Training improved performance on tasks related to prefrontal functioning and had also a significant effect on motor activity in children with ADHD. The results thus suggest that WM training potentially could be of clinical use for ameliorating the symptoms in ADHD.

The ability to retain and manipulate information in WM depends on the prefrontal cortex (Fuster, 1989; Goldman-Rakic, 1987) and underlies several cognitive abilities, including logical reasoning and problem-solving (Engle, Kane, & Tuholski, 1999; Hulme & Roodenrys, 1995; Klingberg, 2000). Furthermore, WM capacity has been regarded as a permanent trait of the individual, closely related to *g*, a proposed measure of general cognitive ability (Engle, Kane, & Tuholski, 1999; Kyllonen & Christal, 1990). Attention deficit hyperactivity disorder (ADHD) is characterized by inattention, impulsivity, and hyperactivity (American Psychiatric Association, 1994). Among the cognitive deficits in this disorder, WM impairment is of central impor-

tance (Barkley, 1997; Kuntsi, Oosterlaan, & Stevenson, 2001; Mariani & Barkley, 1997; Rapport, Chung, Shore, Denney, & Isaacs, 2000; Westerberg, Hirvikoski, Forssberg, & Klingberg, submitted) and has been suggested to be associated with an impaired function of the frontal lobe (Rubia et al., 1999; Schweitzer et al., 2000; Zametkin et al., 1990).

In the present study we investigated whether WM capacity could be improved by training. Furthermore, if impairment of WM is a core deficit in ADHD, this would imply that improvement of WM would decrease the symptoms in ADHD. Previous attempts to improve WM by training have only achieved moderate success. In experiments where subjects perform repeated

WM trials, without adapting the difficulty level, this typically only leads to faster reaction times, but no increase in WM capacity has been reported (Kristofferson, 1972; Phillips & Nettelbeck, 1984). Some success has been achieved in teaching rehearsal strategies to children with learning disabilities (Brown, Campione, Bray, & Wilcox, 1973; Butterfield, Wambold, & Belmont, 1973; Hulme, 1992). There are also case studies of subjects who have learned strategies to retain a large number of digits (Ericsson, Chase, & Faloon, 1980). However, in these studies the strategies were not useful for improving on other WM or reasoning tasks. Therefore, such training does not increase general WM capacity.

Here, we investigated whether WM capacity could be improved by a new type of computerized, cognitive training and whether this would reduce motor activity in children with ADHD. We adopted two key features of a training regime previously used to enhance sensory discrimination and induce cortical plasticity in sensory and motor cortices (Buonomano & Merzenich, 1998; Tallal et al., 1996): (1) training was performed close to the capacity of the individual by using an adaptive staircase method that adjusted difficulty on a trial-by-trial basis; and (2) training was performed at least 20 min per day, 4–6 days a week, for at least 5 weeks.

We designed a computer program based upon these principles in which subjects practiced WM tasks. Fourteen children with ADHD undertook training of a visuo-spatial WM task, a visuo-spatial version of backwards digit-span, and a spatial-verbal WM task. Visual and verbal feedback was implemented in the computer program to increase compliance during the training. We also designed a "placebo" or "low-dose" computer-program, which was similar to the treatment-program, but did not include the two key features: thus difficulty level was not interactively adjusted, and daily training amounted to less than 10 min per day. The study was designed as a double-blind study where children, parents, and the psychologist administrating pre- and posttraining tests were blinded to which version of the computer program the children had practiced and to the difference in expected effect of the two versions.

Subjects performed a battery of cognitive tasks before and after training to evaluate WM capacity and prefrontal functioning. These tasks included Raven's Coloured Progressive Matrices (Raven, 1995), which is a series of complex reasoning tasks thought to measure prefrontal functioning and general cognitive ability. Performance on this task has been estimated to have about $r^2 = .9$ correlation with g (Engle et al., 1999). Impulsivity was estimated using the Stroop task, on which children with ADHD are known to have impaired ability to perform. The Stroop task also relies on activation of the prefrontal cortex (Bench et al., 1993; Pardo, Pardo, Janer, & Raichle, 1990). To get a measure of motor activity level before and after training we used an infrared camera that detected movements of a marker placed on the child's head while the child performed a 15-min continuous performance task on a computer. Motor activity as measured by this method has previously been shown to correlate with behavioral ratings of hyperactivity in children with ADHD (Teicher, Ito, Glod, & Barber, 1996) and to be sensitive to measure the effects of stimulant treatment (Teicher et al., 2000). In a subsequent experiment, we tested whether young adults without ADHD or WM deficits would be able to increase WM capacity by training.

EXPERIMENT 1

METHODS

Subjects
Children participating in the study were between 7 and 15 years of age and diagnosed with ADHD by a pediatrician according to the guidelines of DSM-IV. The treatment group included 1 girl and 6 boys (mean age = 11.0, $SD = 2.0$), and the control group included 2 girls and 5 boys (mean age = 11.4, $SD = 3.0$). Three subjects in the treatment group and 2 subjects in the control group were on medication. Movement analysis could not be performed in 2 subjects in the treatment group due to technical failures. There were no significant differences in age between the groups, or any differences in pretest scores for either of the cognitive task or for head movements. The study was conducted at the Astrid Lindgren Children's Hospital, which is part of the Karolinska Hospital

and approved by the local ethics committee at the Karolinska Hospital.

Tests for Pre- and Posttraining Evaluation

All subjects were tested on five cognitive tasks: (1) Trained version of the visuo-spatial WM task: Circles were presented one at a time in a four-by-four grid. After a delay the subjects indicated the positions of the circles. The number of circles in the sequence was successively increased until the subject missed two trials in a row. The score was the maximum number of circles remembered, with half a point given for one correct trial out of two. (2) Span board: Ten blocks were arranged in an irregular pattern in front of the subject. The testing psychologist pointed to a sequence of blocks and the subject then pointed to the same blocks in the same order (forward version) or in the reverse order (backward version). Scoring was performed as for the trained version of the visuo-spatial WM task. (3) Stroop task: Words describing colors were printed with ink in a color that was incongruent with the word, that is, ''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: A series of nonverbal reasoning tasks described in detail elsewhere (Raven, 1995). This test is often used for estimating general nonverbal mental ability. (5) Choice reaction time task. A computer was used for displaying visual stimuli. A separate serial-response box (Psychology Software Tools, Pittsburg) was used for collecting responses. Yellow circles would appear in one out of two possible locations on the screen, to the left or to the right. The task was to press a button when a yellow circle appeared. The yellow circle was preceded by a warning cue consisting of a grey circle that appeared 1–4 s before the yellow circle. Each subject was first tested on a simple-reaction time version with only one possible stimulus location (first circles appearing on the left location, with responses made by the left hand, then repeated on the right hand side). Then subjects were tested on a choice reaction time task with warning cues appearing simultaneously to the right and left, and yellow circles appearing randomly either in the left or the right location. The dependent measures were (1) reaction time for one-choice trials; (2) increase in reaction time for two-choice as compared with one-choice trials; and (3) variance of reaction times.

The measurement of head movements has been described in previous publications (Teicher et al., 1996). An infrared motion analysis system (OPTAx Systems, Burlington, MA) recorded the movements of a small reflective marker attached to the back of the head of the child. A movement was defined to begin when the marker moved 1.0 mm or more from its most recent resting location. The number of movements was recorded during a 15-min period when the child was performing a version of a continuous performance task. In this task subjects were asked to respond to a target and withhold response to nontargets, with no requirement of holding any information in WM. Stimuli were presented every 2.0 s, and 50% of stimuli were targets.

Computerized Training Program

Four subtests were presented during each training session: (1) A visuo-spatial WM task where circles were presented one at a time in a four-by-four grid (same task as the one used for testing). (2) Backwards digit-span. A keyboard with numbers was shown and digits read aloud. The subject then marked the digits, but in the reverse order. (3) Letter-span task. Letters were read aloud one at a time. The subject had to remember the identity and order of the letters. A row of lamps was then visible and a flashing lamp cued the subject as to which letter should be reported back, for example, if lamp no. 3 was lit, the subject should report the third letter that they previously heard. (4) Choice reaction time task. This task was not a WM task but a mixture of a reaction-time task and a go/no-go task, and was included based on the evidence from our laboratory that children with ADHD are impaired on similar tasks. Two grey circles were presented on the screen (horizontally oriented in the two-choice condition). Subjects were told to press a spatially congruent key when one of the circles became green, and to withhold responding when one of the circles became red.

For all WM tasks, difficulty was adjusted by changing the number of stimuli to be remembered. Subjects completed 30 trials on each WM task every day, and the daily training time was approximately 25 min. Time between test and retest was 5–6 weeks during which the subjects trained 24.3 (2.2 *SEM*) days. For two types of tasks: the visuo-spatial WM task and the digit-span task, subjects proceeded to train a second and more demanding version after 10–18 days of training on the first version. In this second version a visual distraction occurred during the delay in the WM tasks.

The placebo/low-dose version included 10 trails per task, with 2 stimuli to remember in the visuo-spatial and digit span task, and 3 stimuli in the letter-span task. The placebo version was intended to control for the effect of taking the evaluation test repeatedly, for the spontaneous improvement that could occur over the training period of 5–6 weeks, and for possible bias for the psychologist administering the tasks as well as for effects of expectancy from the children. The placebo-version did not, however, control for the total amount of time in front of a computer.

RESULTS

Test-retest changes in the group of subjects undertaking the treatment program (treatment group) was compared to the test-retest changes in the group of subjects using the placebo program (control group). This comparison demonstrated a significant treatment effect for the practiced visuo-spatial WM task (Table 1) as well as for the Span board task, a nonpracticed visuo-spatial WM task (Fig. 1a; Table 1). In the Span board task, subjects remembered the position of small cubes arranged in a pseudo- random order. While the task tested visuo-spatial WM, it differed in several ways from the practiced visuo-spatial WM task: it was not computerized, but involved three-dimensional stimuli, had a different spatial arrangement of stimuli, and involved different modes of stimulus presentation and response. Substantial improvements were evident for all children on the Span board task (Fig. 1a), and group differences were significant ($P < .0001$; performance on the backward and forward versions were averaged). Significant improvement on Raven's Progressive Matrices was also evident (Table 1). Again, all children in the treatment group improved (Fig. 1b). Significant group difference was found also for the Stroop task (Table 1). However, only weak and inconsistent effects were seen in a choice reaction time task (Table 1).

The number of head movements was significantly reduced in the treatment group compared to the control group (Table 1, Fig. 1c). Again, this effect was evident in all subjects in the treatment group (Fig. 1c). The number of head movements during retest in the control group was about 6% higher than during the first testing. This is consistent with previous data on test-retest changes after administration of pharmacological placebo, where an increase of about 8% was found on the second testing (Teicher et al., 2000; Teicher, personal communication). The reduction of head movements in the treatment group was 74% (*SEM* 7). In comparison, a probe dose of methylphenidate (approximately 0.4 mg/kg) reduced the number of head movements by 62% (Teicher et al., 2000).

Test-retest and the treatment-control comparisons were significant for four tests: trained WM, Span board, Raven's Progressive Matrices, Stroop accuracy, and number of head movements (Table 1). Correlation analysis on the test-retest differences showed that improvement in the two WM tasks were significantly correlated, as was the correlation between improvement on Raven's Progressive Matrices and the trained WM task (Table 2). The reduction in head movements was highly correlated with improvements on both the trained WM task ($r^2 = .73$) and Raven's Progressive Matrices ($r^2 = .74$).

EXPERIMENT 2

METHODS

Subjects
Four male, healthy volunteer subjects (AP, DE, DH, IK), aged 23, 29, 20 and 22 years, who were all university students, participated in the training. The subjects had no history of psychiatric or neurological disease. The local ethics committee at the Karolinska Hospital approved the study.

Tests for Pre- and Posttraining Evaluation
Before and after training, all subjects were tested on similar tasks as described in Experiment 1: (1) Trained version of the visuo-spatial WM task; (2) Span board task; (3) Stroop task; (4) Raven's Advanced Progressive Matrices; and (5) Choice reaction time task. Posttraining results on the choice reaction time task was lost for two of the subjects. In contrast to Experiment 1, the Advanced Progressive Matrices were given (Raven, 1990). Eighteen problems (even numbered problems) were given before testing, and a new set of 18 problems (odd numbered problems) was given afterwards. No measurement of head movement was made.

Subjects undertook 5 weeks of training on the same computerized training program as described in Experiment 1, with an average of 26 days of training.

RESULTS

During training, performance improved gradually on all trained tasks, with an increased amount of information kept in WM and decreased reaction-times (Fig. 2). On the cognitive testing performed before and after training, improvement was evident for all subjects and all tasks (Table 3). Adult

Table 1. Task Performance in Children With ADHD in the Treatment Group, Before and After Training.

	Control		Treatment		Test-retest in treatment group[a]	Group difference[b]
	Before Mean (*SEM*)	After Mean (*SEM*)	Before Mean (*SEM*)	After Mean (*SEM*)		
Trained visuo-spatial WM	5.0 (0.22)	4.79 (0.21)	4.71 (0.21)	6.43 (0.41)	$P = .0007$	$P = .0006$
Span board[c]	4.54 (0.46)	4.93 (0.28)	4.36 (0.12)	6.32 (0.25)	$P = .0001$	$P = .001$
Stroop task						
Accuracy (max 60)	56.3 (0.8)	55.1 (2.6)	55.4 (1.2)	59.4 (0.3)	$P = .03$	$P = .02$
Time for completion (s)	80.3 (6.7)	86.3 (15.1)	101 (7.9)	90.9 (7.2)	$P = .17$	$P = .12$
Raven's progressive matrices	28.7 (0.8)	29.3 (1.0)	26.4 (1.2)	33.3 (1.6)	$P = .001$	$P = .001$
Choice reaction time task						
RT Latency (ms)	314 (18)	342 (30)	282 (22)	296 (31)	$P = .20$	$P = .27$
Two – One choice (ms)	91 (19)	79 (17)	146 (24)	71 (14)	$P = .07$	$P = .05$
RT standard deviation	128 (31)	117 (20)	106 (24)	92 (17)	$P = .18$	$P = .49$
Number of head movements	1496 (579)	1881 (616)	1001 (269)	315 (148)	$P = .002$	$P = .00008$

Note. [a]Test-retest differences (before and after training) in the treatment group (one-tailed *t*-test).
[b]Intra-individual test-retest differences were compared between the treatment group and the placebo group (one-tailed *t*-test).
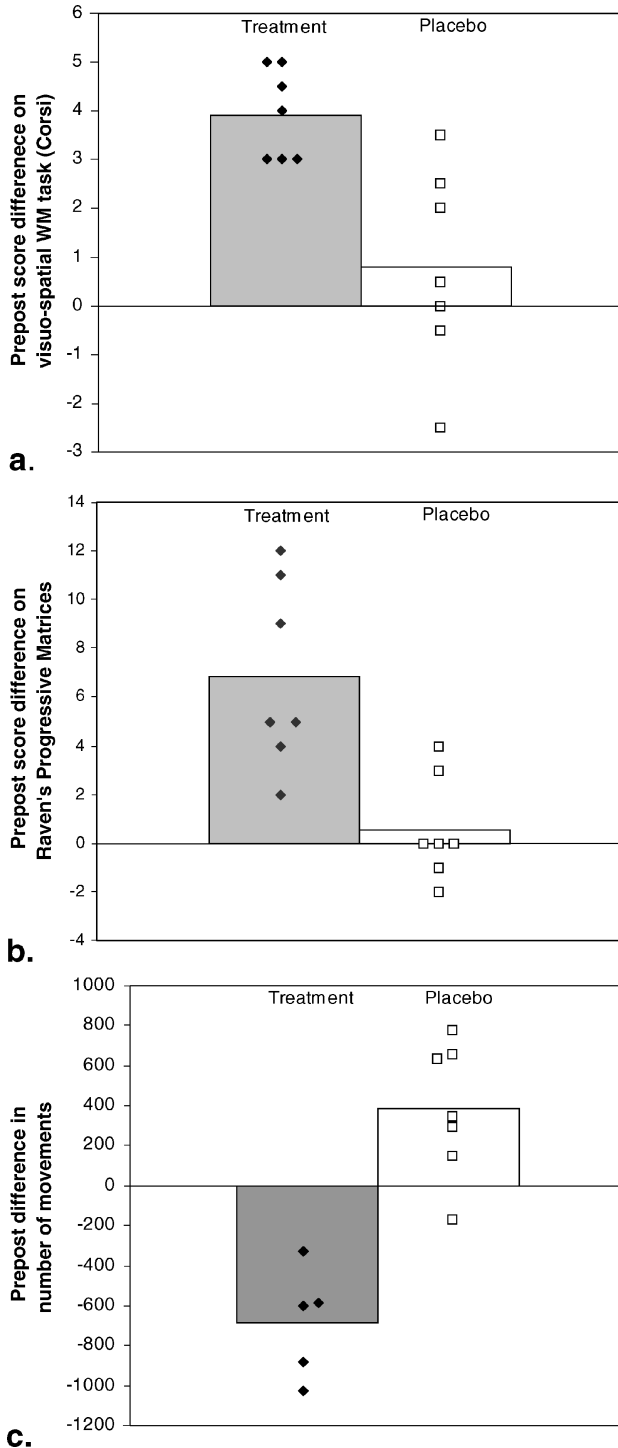[c]Mean improvement for both backward and forward version.

Fig. 1. Test-retest differences in the treatment and placebo group. Bars show mean values for each group. Markers represent individuals in each group. The group differences were significant for each of the three tasks (Table 1). (a) Span board task, a nontrained visuo-spatial WM task; (b) Raven's Progressive Matrices; (c) number of head movements.

Table 2. Correlations (r^2) Between Test-Retest Differences in the Treatment Group.

	Trained WM	Corsi block	RPM	Head movements
Span board	.85**			
RPM	0.74*	0.55		
Head movements	0.73	0.56	0.74	
Stroop accuracy	0.42	0.59	−0.22	0.20

Note. RPM = Raven's Progressive Matrices.
  *$P < 0.05$; **$P < 0.01$.

test-retest values were compared to those of the control group in Experiment 1. Alternatively, the test-retest scores for the adults could have been compared with an adult control group. However, because it is the change between test and retest that is compared, the differences in pretest scores are subtracted away. The test-retest improvement of the control group in Experiment 1 is also consistent with previous reports of test-retest improvements in adults on spatial WM (Lowe & Rabbitt, 1998) and the Stroop task (Salinsky, Storzbach, Dodrill, & Binder, 2001). Compared to the placebo-group in Experiment 1, the improvement was significant for the trained visuo-spatial WM task, Span board, Stroop time and Raven's Progressive Matrices ($P < .05$). However, all subjects were correct on the highest level included on the visuo-spatial WM tasks (level 9) and 2 achieved highest score on the Span board (level 8, both on the forward and the backward version), and 2 subjects achieved maximum score on the Stroop task (100). The differences between the trained group and the control group in Experiment 1 was thus probably underestimated due to ceiling effects in the adult group. On Raven's Progressive Matrices no subject achieved maximum score and the improvement was 24.5%, which is similar to the 26.1% improvement for

Table 3. Task Performance in Four Healty Adults, Before and After Training.

	Before Mean (*SEM*)	After Mean (*SEM*)	Improvement by treatment[a]	Group difference[b]
Trained visuo-spatial WM	7.12 (0.62)	9.0 (0.00)	$P = .03$	$P = .005$
Span board[c]	5.6 (0.39)	7.25 (0.48)	$P = .02$	$P = .02$
Forwards (items)	5.62 (0.71)	7.0 (0.81)	$P = .02$	$P = .02$
Backwards (items)	5.62 (0.24)	7.50 (0.35)	$P = .01$	$P = .06$
Stroop task				
Accuracy (max 100)	98.25 (0.85)	99.5 (0.29)	$P = .10$	$P = .24$
Time for completion (s)	109.7 (7.4)	86.7 (4.1)	$P = .01$	$P = .02$
Raven's progressive matrices				
(max 18)	12.25 (0.25)	15.25 (0.85)	$P = .01$	$P = .04$
Choice reaction time task				
RT Latency (ms)	248 (16)	220 (5)	$P = .12$	$P = .06$
Two – One choice (ms)	60 (25)	44 (3)	$P = .30$	$P = .43$
RT standard deviation	54 (17)	32 (10)	$P = .10$	$P = .42$

Note. [a]Paired *t*-tests of test-retest differences (before and after training) in the treatment group (one-tailed *t*-test).
  [b]Intra-individual test-retest differences were compared between the treatment group and the placebo group in Experiment 1 (one-tailed *t*-test).
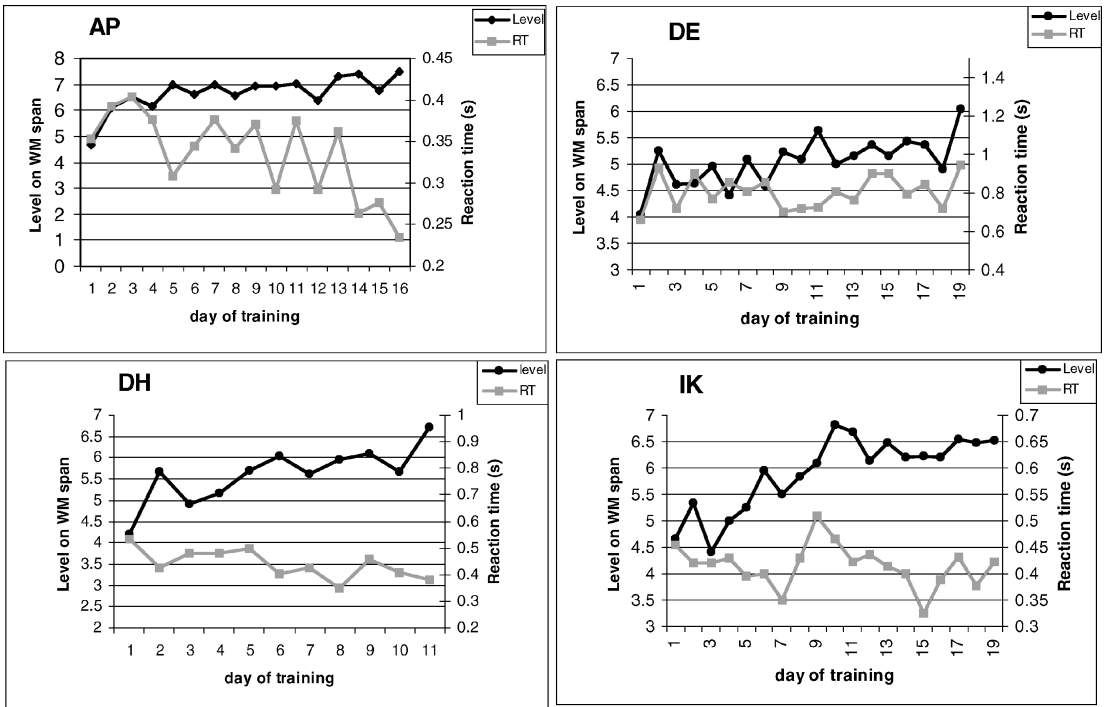 [c]Mean improvement for both backward and forward version.

Fig. 2. Improvement during training of the visuo-spatial WM task in four adult subjects. Each graph shows data from a single individual.

children in the treatment group in Experiment 1. Raven's Progressive Matrices is known as a highly reliable test where test-retest improvement is typically 0–8% (Raven, 1990). With test-retest intervals as short as 2 weeks it has been shown to be 9% (Costenbader & Ngari, 2001). With longer retest intervals the improvement is less, and in Experiment 1 the test-retest improvement after 5 weeks was 2%.

DISCUSSION

The present study showed that intensive and adaptive, computerized WM training gradually increased the amount of information that the subjects could keep in WM (Tables 1 and 3, Figs. 1 and 2). The improved performance occurred over weeks of training, and is in this respect similar to the slow acquisition of a perceptual skill or a motor skill (Karni et al., 1995; Recanzone, Schreiner, & Merzenich, 1993; Tallal et al., 1996). Furthermore, the improvement from training was evident

both for a group of children with ADHD (Experiment 1), as well as for adult subjects without ADHD (Experiment 2). This shows that an initial deficit in WM was not necessary for improvement to occur.

Increased performance was seen for both trained and nontrained visuo-spatial WM tasks, showing that the training effect generalized to other settings (Tables 1 and 3). Only inconsistent effects where seen for a choice reaction time task (Table 1), which is in agreement with a previous observation that cognitive tasks are more accessible for improvement from training than are attentional tasks (Sohlberg, McLaughlin, Pavese, Heidrich, & Posner, 2000). A significant training effect was also seen for the reasoning task (Tables 1 and 3, Fig. 1b). Although the improvement on the Span board task was evidence some transfer of training, it is difficult to evaluate the commonalities of the two tasks, and to exclude that the subjects were trained towards improvement on the Span board task. However, the improvement on the reasoning task is a clear

evidence of that the training effect generalized to nonpracticed tasks, since the training did not include any problem solving or reasoning exercises at all. This is true also for the enhanced performance on the Stroop task. The improvement in reasoning ability is likely due to the fact that complex reasoning depends on WM, or more precisely, that the trained WM tasks and the reasoning task rely on the same cortical areas. A common factor underlying the WM tasks, Raven's Progressive Matrices and the Stroop task is that they all depend on the prefrontal cortex (Bench et al., 1993; Duncan & Owen, 2000; Klingberg, Forssberg, & Westerberg, 2002; Pardo et al., 1990; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997). The association between the reasoning task and the WM tasks is further substantiated by the significant correlation between improvement on the visuo-spatial WM task and improvement on Raven's Progressive Matrices (Table 2).

The second main finding was that training on WM tasks significantly reduced the number of head movements (Table 1, Fig. 1c), with improvements on WM being correlated with the reduction in movements (Table 2). The measurement of head movements was an attempt to get an objective measurement of the children's motor activity. This measure of motor activity has previously been shown to correlate with behavioral ratings (Teicher et al., 1996). An obvious drawback of the test is that it only measures behavior during a short period of time, and in a laboratory situation. It would therefore ideally be complemented with additional measurement of hyperactivity. Nevertheless, the decrease in head movements gives a clue to the question of how the cognitive deficits, the impulsivity and the motor symptoms in ADHD are related to each other, and whether some symptoms are more fundamental than others (Barkley, 1997). The present results, where increased WM capacity resulted in decreased movements, suggest that there could be a causal relationship between cognitive functioning and motor behavior.

The fact that WM capacity including reasoning ability, could be substantially affected by training deserves some comment since this seems to challenge the notion that WM capacity reflects a fixed property of the individual (Engle et al.,

1999; Kyllonen & Christal, 1990). The results also seem to contradict the notion that tasks such as Raven's Progressive Matrices measure a fixed cognitive ability (Jensen, 1998). However, only one reasoning test was given, and to test the hypothesis about how experience affects cognitive ability, a wider battery of tasks would be needed.

Future research will have to investigate to what extent the training also effects everyday life for children with ADHD, and document the durability of the training effects. However, here we could show that training significantly improved performance on tasks related to cognitive function, inhibition (as measured by the Stroop task), as well as motor activity, which suggests that WM training could be of interest for future clinical use in children with ADHD.

ACKNOWLEDGMENTS

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

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

Bench, C.J., Frith, C.D., Grasby, P.M., Friston, K.J., Paulesu, E., Frackowiak, R.S., & Dolan, R.J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia, 31,* 907–922.

Brown, A.L., Campione, J.C., Bray, N.W., & Wilcox, B.L. (1973). Keeping track of changing variables: Effects of rehearsal training and rehearsal prevention in normal and retarded adolescents. *Journal of Experimental Psychology, 101,* 123–131.

Buonomano, D.V., & Merzenich, M.M. (1998). Cortical plasticity: From synapses to maps. *Annual Reviews of Neuroscience, 21,* 149–186.

Butterfield, E.C., Wambold, C., & Belmont, J.M. (1973). On the theory and practice of improving short-term memory. *American Journal of Mental Deficiency, 77,* 654–669.

Costenbader, V., & Ngari, S.M. (2001). A Kenya Standardization of the Raven's Coloured Progressive Matrices. *School Psychology International, 22,* 258–268.

Duncan, J., & Owen, A.M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands [In Process Citation]. *Trends Neuroscience, 23,* 475–483.

Engle, W.R., Kane, J.M., & Tuholski, S.W. (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 A. Myake & P. Shah (Eds.), *Models of working memory* (pp. 102–134). Cambridge: Cambridge University Press.

Ericsson, K.A., Chase, W.G., & Faloon, S. (1980). Acquisition of a memory skill. *Science, 208,* 1181–1182.

Fuster, J.M. (1989). *The prefrontal cortex.* New York: Raven Press.

Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In V.B. Mountcastle (Ed.), *Handbook of physiology* (pp. 373–417). Bethesda, MD: American Physiological Society.

Hulme, C. (1992). *Working memory and severe learning difficulties.* Hove: Lawrence Erlbaum Associates.

Hulme, C., & Roodenrys, S. (1995). Practitioners review: Verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry, 36,* 373–398.

Jensen, A.R. (1998). *The g factor: The science of mental ability.* Westport, Conneticut: Praeger.

Karni, A., Meyer, G., Jezzard, P., Adams, M.M., Turner, R., & Ungerleider, L.G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377,* 155–158.

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

Klingberg, T. (2000). Limitations in information processing in the human brain: Neuroimaging of dual task performance and working memory tasks. *Progress in Brain Research, 126,* 95–102.

Kristofferson, M.W. (1972). Effects of practice on character-classification performance. *Canadian Journal of Psychology, 26,* 54–60.

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

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14,* 389–433.

Lowe, C., & Rabbitt, P. (1998). Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: Theoretical and practical issues. Cambridge Neuropsychological Test Automated Battery. International Study of Post-Operative Cognitive Dysfunction. *Neuropsychologia, 36,* 915–923.

Mariani, M.A., & Barkley, R.A. (1997). Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. *Developmental Psychology, 13,* 111–129.

Pardo, J.V., Pardo, P.J., Janer, K.W., & Raichle, M.E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences of the United States of America, 87,* 256–259.

Phillips, C.J., & Nettelbeck, T. (1984). Effects of practice on recognition memory of mildly mentally retarded adults. *American Journal of Mental Deficiency, 88,* 678–687.

Prabhakaran, V., Smith, J.A., Desmond, J.E., Glover, G.H., & Gabrieli, J.D. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33,* 43–63.

Rapport, M.D., Chung, K.M., Shore, G., Denney, C.B., & Isaacs, P. (2000). Upgrading the science and technology of assessment and diagnosis: Laboratory and clinic-based assessment of children with ADHD. *Journal of Clinical Child Psychology, 29,* 555–568.

Raven, J.C. (1990). *Advanced Progressive Matrices. Set II.* Oxford, United Kingdom: Oxford Psychologists Press.

Raven, J.C. (1995). *Coloured Progressive Matrices.* Oxford, United Kingdom: Oxford Psychologists Press.

Recanzone, G.H., Schreiner, C.E., & Merzenich, M.M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience, 13,* 87–103.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S.C., Simmons, A., & Bullmore, E.T. (1999). Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: A study with functional MRI. *American Journal of Psychiatry, 156,* 891–896.

Salinsky, M.C., Storzbach, D., Dodrill, C.B., & Binder, L.M. (2001). Test-retest bias, reliability, and regression equations for neuropsychological measures repeated over a 12–16-week period. *Journal of the International Neuropsychology Society, 7,* 597–605.

Schweitzer, J.B., Faber, T.L., Grafton, S.T., Tune, L.E., Hoffman, J.M., & Kilts, C.D. (2000). Alterations in the functional anatomy of working memory in adult attention deficit hyperactivity disorder. *American Journal of Psychiatry, 157,* 278–280.

Sohlberg, M.M., McLaughlin, K.A., Pavese, A., Heidrich, A., & Posner, M.I. (2000). Evaluation of attention process training and brain injury education in persons with acquired brain injury. *Journal of Clinical and Experimental Neuropsychology, 22,* 656–676.

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S.S., Schreiner, C., Jenkins, W.M., & Merzenich, M.M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271,* 81–84.

Teicher, M.H., Anderson, C.M., Polcari, A., Glod, C.A., Maas, L.C., & Renshaw, P.F. (2000). Functional deficits in basal ganglia of children with attention-deficit/hyperactivity disorder shown with functional magnetic resonance imaging relaxometry. *Nature Medicine, 6,* 470–473.

Teicher, M.H., Ito, Y., Glod, C.A., & Barber, N.I. (1996). Objective measurement of hyperactivity and attentional problems in ADHD. *Journal of American Academy for Child and Adolescent Psychiatry, 35,* 334–342.

Westerberg, H., Hirvikoski, T., Forssberg, H., & Klingberg, T. (submitted). Visuo-spatial working memory: A sensitive measurement of cognitive deficits in ADHD.

Zametkin, A.J., Nordahl, T.E., Gross, M., King, A.C., Semple, W.E., Rumsey, J., Hamburger, S., & Cohen, R.M. (1990). Cerebral glucose metabolism in adults with hyperactivity of childhood onset. *New England Journal of Medicine, 323,* 1361–1366.