Working memory training improves arithmetic skills and verbal working memory capacity in children with ADHD
Children with ADHD diagnosis often display working memory deficits, as well as reading and mathematical disabilities. Previous studies have demonstrated that computerized working memory training (WMT) is a promising intervention. The present study aimed at exploring the effects of WMT on working memory, scholastic skills and behavioral symptoms in children with ADHD. Thirty-two children, aged 6 to 11, were randomized to WMT or a control condition. WMT consisted of nine tasks taxing working memory with adaptive difficulty level. All children trained in their homes, with their parents acting as supervisors. Children who completed more than 20 days of training in 5-8 weeks (8 in the WMT condition and 13 in the control condition) were considered compliers. Assessments were conducted before and after intervention. Results indicated that WMT lead to significant gains of verbal working memory and arithmetic skills. More research is needed to further investigate the effects of WMT.

Computerized working memory training (WMT) is a relatively new intervention and the subject of a growing field of research. WMT software is available in Swedish schools and hospitals and is used by both children and adults. The present study is a randomized controlled trial of WMT in children with ADHD. Effects of WMT and reading training (RT) on working memory capacity, scholastic skills and behavioral symptoms were compared.

Attention Deficit Hyperactivity Disorder (ADHD) is a developmental mental disorder characterized by inattention and/or hyperactivity/impulsivity (American Psychiatric Association, 1994). At least six out of nine diagnostic criteria regarding attention or six out of nine diagnostic criteria regarding hyperactivity/impulsivity need to be met in order to be diagnosed. Three subtypes are thereby stipulated; ADHD-predominantly hyperactivity-impulsive type (ADHD-H), ADHD-predominantly inattentive type (ADHD-I) and ADHD-combined type (ADHD-C). Hyperkinetic disorder, a term used by The World Health Organization (1993) contains similar criteria as the DSM-IV diagnosis. In the review of earlier research below, Hyperkinetic disorder will not be differentiated from ADHD.

A recent meta-analysis of the prevalence of ADHD (Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007) concluded that around 5% of children and adolescents worldwide fulfill diagnostic criteria for ADHD. A significant association between gender and prevalence was found in the same study, revealing that boys were more than twice as likely to be diagnosed with ADHD as girls.
Even though ADHD-symptoms seem to decrease in adulthood, problems persist in some degree in a substantial part of cases (Faraone, Biederman, & Mick, 2006). The distribution of subtypes have varied in studies, but in one study investigating different subtyping methods (Rowland, et al., 2008), ADHD-H was the least common subtype, and either ADHD-C or ADHD-I were most common depending on subtyping procedure.

ADHD has often been associated with the presence of one or more comorbid disorders and scholastic disabilities. Correlations have been found between ADHD and for example oppositional defiant behaviors, autistic traits, motor coordination problems, anxiety (Gillberg, et al., 2004), reading problems (Rommelse, et al., 2009), mathematical disabilities (Faraone, et al., 1993) and Tourette’s syndrome (Kadesjö & Gillberg, 2000).

ADHD theories
Several different theories of the developmental paths and mechanisms of ADHD have been suggested. Barkley’s (1997) influential model proposed that behavioral inhibition was the primary deficit of ADHD. Behavioral inhibition was considered to be fundamental since it was thought to create a delay in responding and thus a time window in which executive functions could become activated. The theory specified four separable executive functions based on prior research: verbal working memory, nonverbal working memory, the self-regulation of affect/motivation/arousal, and reconstitution (analyzing behavior and creating novel behavioral responses). The primary impairment in behavioral inhibition was hypothesized to cause secondary impairments in executive functioning which in turn resulted in failures in motor control, fluency and syntax.

Another core deficit of ADHD was later proposed by Sagvolden, Aase, Zeiner and Berger (1998). Based on the fact that the effect of reinforcement decreases when the time window between response and reinforcement grows, the authors hypothesized that, ADHD children would be more sensitive to reinforcement in close proximity and less sensitive to distal reinforcement compared to children without the diagnosis. The altered reinforcement hypothesis further suggested that the often observed hyperactivity and problems with sustained attention in the ADHD population could be explained by the sensitivity to delay of reinforcement.

A later proposed theory (Sonuga-Barke, 2002) suggested a synthesis between the theories of executive dysfunction and delay sensitivity, comprising a dual pathway model of ADHD-C. The reason for joining the two processes in one model was that they both seemed to relate to ADHD even though they did not correlate with each other. Thus the dual pathway model stated that symptoms labeled as ADHD could be the result of either of the two independent developmental paths. The first pathway described how inhibitory dysfunction leads to cognitive and behavioral dysregulation. The cognitive dysregulation then resulted in poor quality task engagement, while the behavioral dysregulation resulted in the ADHD symptoms. The second pathway to ADHD stated that the sensitivity to delay (shortened delay gradient) in combination with contextual demands of waiting and delay resulted in repeated failures. This
resulted in an acquired generalized delay aversion through mechanisms of associative conditioning. The delay aversion in turn led to the ADHD symptoms.

Another theory is the cognitive-energetic model, put forth by Sergeant (2000), which also included components of delay aversion and behavioral inhibition. The model stipulated that the problems associated with ADHD were evident on three different levels: the executive system (for example behavioral inhibition and working memory), state factors and the computational mechanisms of attention. The state factors, or energetic pools as they were also called, consisted of effort, arousal and activation. The cognitive-energetic model suggested that dysfunction in the state factors might be underlying the behavioral inhibition evident in ADHD.

One implication of the behavioral inhibition deficit theory was that executive dysfunction ought to be found, in some degree, in all cases of ADHD. A comprehensive meta-analysis (Willcutt, Doyle, Nigg, Faraone, & Penninger, 2005) recently came to the conclusion that, although executive dysfunction is significantly related to ADHD, it is neither necessary, nor sufficient to cause ADHD. Instead, executive dysfunction was shown to be one of several important deficits characterizing ADHD. The executive functions that were found to have the strongest effects were response inhibition, vigilance, working memory, and some measures of planning. Another recent meta-analysis (Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005) also investigated working memory impairments in ADHD and found strong impairments in spatial working memory and moderate impairments in verbal working memory. Consequently, a conclusion based on the theoretical models of ADHD and the above mentioned meta-analysis concerning executive dysfunction in ADHD, is that working memory deficits are common in ADHD.

ADHD treatment
The most common treatment for ADHD has been pharmacological, such as methylphenidate (Bedard, Jain, Hogg-Johnson, & Tannock, 2007) or atomoxetine (Cheng, Chen, Ko, & Ng, 2007). Different forms of psychosocial, primarily behavioral, treatments have also been employed in both school and home environment. The effects of behavioral interventions on ADHD have been a subject of controversy for several years. Fabiano et al. (2009) performed a meta-analysis including studies with different designs (single subject, pre-post, between group and within-subject) and concluded that behavioral interventions in different settings had a positive effect on ADHD symptoms. Examples of included behavioral interventions were parent training with emphasis on social learning principles, behavior modification techniques and behavioral classroom interventions. A meta-analysis by Van der Oord, Prins, Osterlaan, and Emmelkamp (2008) compared the effects of different forms of cognitive and/or behavioral interventions with the effects of methylphenidate. Effects were found on ADHD-symptoms in school-aged children for both treatments but not on academic functioning. However the psychosocial treatments did not add to the effects of methylphenidate.

Another meta-analysis (Schachter, Pham, King, Langford, & Moher, 2001) concluded that the use of methylphenidate in some cases caused loss of appetite, insomnia, and to a lesser extent stomach ache, headache and dizziness. Effects of long-term use beyond 14 months, in randomized controlled conditions, have not been sufficiently investigated.
Corcoran and Dattalo (2006) have examined the effects of parent-involvement in treatment of ADHD. According to the results of their meta-analysis, including studies of cognitive-behavioral interventions, parent-involvement was associated with a weak positive effect on ADHD and externalizing problems. Effects in the moderate range could be seen on outcome measures such as internalizing symptoms and academic performance, although the results on the later measure were described as tentative.

**Working memory**

Probably the first mention of working memory was by Miller, Galanter and Pribram (1960). However the most influential model of working memory is the multi-component model presented by Baddeley and Hitch (1974). The model comprised of three components; the central executive and its two slave systems the phonological loop and the visuo-spatial sketchpad. Later a third slave system, the episodic buffer, was added to the model (Baddeley, 2000). Figure 1 gives an overview of the revised multi-component model of working memory. Despite the fact that the original model is quite simple and over thirty years old it has continued to develop and to be a subject for research and debate (Baddeley, 2003). There are several additional models of working memory. For a comprehensive review on different models of working memory see Miyake and Shah (1999).

Baddeley (1992) defined working memory as follows: “The term working memory refers to a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning and reasoning” (p. 556).

![Diagram of the multi-component working memory model](image)

**Figure 1.** The multi-component working memory model. The grey area represents crystallized cognitive systems, and the white areas represent fluid cognitive systems (Baddeley, 2000).
The phonological loop has been described as consisting of two subsystems. The first system is the phonological store which is capable of holding acoustic information for about two seconds. The second system is the articulatory control process which has two functions. It repeats information in the phonological store subvocally to maintain the information for longer periods of time. The second function of the articulatory control process is to transform visually presented nameable information (such as letters or words) to enable subvocal repetition (Baddeley, 1992).

The visuospatial sketchpad has been assumed to serve the same function as the phonological loop, but regarding visual and spatial stimuli and information. Although debated, at least two separate components seem to exist: one component dealing with the spatial location of objects and the other component dealing with visual representation of objects. However the research on the visuospatial sketchpad is less extensive than on the phonological loop and yet other components have been suggested in the visuospatial subsystem (Baddeley, 2007).

The central executive, although being the most important component of working memory, is also the least understood (Baddeley, 2003; Baddeley, 2007). The attentional control of behavior is however a key aspect of the central executive (Baddeley, 1992). Dividing and focusing attention are the two most investigated subcomponents of current attentional theories and seem likely to be linked to the central executive. The capacity to switch attention is probably not explained solely by one executive function (Baddeley, 2007).

The episodic buffer was introduced to explain the connection between working memory and episodic long term memory. In contrast to the other slave systems it is not associated with a specific kind of stimuli. The episodic buffer is assumed to be a limited-capacity store capable of integrating information from a variety of sources. It is assumed to feed and retrieve information from episodic long term memory. The information is integrated as episodes that can be extended over space and time (Baddeley, 2000).

The existence of a link between working memory and episodic memory, such as the episodic buffer, could imply that primary deficits in attention processes and working memory would lead to secondary impairments in areas such as verbal learning and episodic memory. Baddeley (2000) was not the first to have suggested such a connection. Atkinson and Shiffring’s (1969) influential model of memory also contained a connection between short-term memory and long-term memory. The model stipulated three different memory stores: the sensory registry, the short-term store, and the long-term store. The sensory registry was believed to be able to store information for milliseconds, the short-term memory for up to 30 seconds, and the long-term memory permanently. The model demonstrated the importance of short-term memory for processes such as encoding and retrieval of memories in long-term memory. Mealer, Morgan and Luscomb (1996) revealed that a group of boys diagnosed with ADHD performed worse than controls on some neuropsychological test measuring different aspects of the memory system. The authors concluded that although problems with initial processing and short-term storage were apparent, no deficits in long-term
memory were evident. The children had no difficulties retrieving information from the long-term store.

**Difficulties associated with working memory**

Several studies have investigated the connection between working memory and arithmetic skills. Results have consistently showed that working memory is a strong predictor of mathematical performance (DeStefano & Lefevre, 2004; Passolunghi, Vercelloni, & Schadee, 2007; De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière, 2009). In a large meta-analysis comprising of data from 77 studies with over 6000 participants Daneman and Merikle (1996) concluded that working memory is also a strong predictor of reading comprehension.

Gathercole and Alloway (2008) have described how children with poor working memory struggle when developing reading and mathematical skills. There are several key aspects involved when learning to read. Spelling patterns of individual words must be learned. Furthermore mappings between sounds and individual letters and letter combinations need to be mastered. Particularly the learning of these connections develops slowly in children with poor working memory. In activities designed to attain basic literacy, children with poor working memory often fail due to high demands on working memory. When a sentence or a fragment of text is to be understood the child needs to keep all the information in working memory long enough to interpret what has been read and at the same time understand each word. Children with poor working memory therefore have even more difficulties in interpreting the meaning of a text since they often have difficulties with both decoding individual words and keeping the whole sentence in memory. In mathematics, particularly addition, subtraction, division and multiplication propose difficulties for children with poor working memory. Working memory overload often occurs in exercises designed to develop knowledge of basic number rules. This results in frequent errors and slows down the acquisition of basic skills. Arithmetic puts high demands on working memory as well as retrieval and application of these rules that are likely to not have been learned. Children with poor working memory therefore often use strategies such as finger counting which puts higher demands on working memory than retrieving rules.

One of the most commonly described problems for children with poor working memory is the ability to follow instructions (Gathercole & Alloway, 2008). Many tasks in school involve instructions in several steps. For example the instruction to color the first flower on the paper red, then color every third flower red and all other flowers blue. Children with poor working memory often fail to remember all steps of such instructions and fail to complete the task. This in combination with poor academic skills lead to difficulties in monitoring the quality of work performed.

**Computerized working memory training**

In recent years several studies on WMT have been published, although only three of these studies examined effects of such training on children diagnosed with ADHD. A first preliminary, double-blind, non-randomized study (Klingberg, Forssberg, & Westerberg, 2002) showed promising results. The children participating in the study were between 7 and 15 years of age. Children placed in the treatment group (n=7) trained with a computerized training program consisting of four different tasks: three
tasks taxing on different aspects of working memory (span-board, letter-span and backwards digit-span) and a choice reaction time task (a mixture between a reaction-time task and a go/no-go task). The children in the treatment group trained for approximately 25 minutes per day on a total of 24.3 occasions on average distributed over five to six weeks. It is unclear whether the children trained at home, in school or at another location. A key feature of the computer program given to the children in the training group was that the difficulty level (number of presented items) of the training constantly and automatically changed based on performance on prior training rounds. Therefore the training was always performed in proximity of the individual child’s maximum working memory capacity. The control group (n=7) were given a very similar computer program without the key feature of adjusting difficulty levels. In the placebo version of the program, training remained on the initial low level with two or three stimuli to remember on the different tasks during the entire training period. A flaw of this early study was that the children in the control group’s daily training (less than 10 minutes) didn’t nearly match the time the children in the treatment group spent on training. However the results indicated that difficulty-adapted training could be used to improve visuo-spatial working memory (span-board), non-verbal complex reasoning (Raven’s Progressive Matrices) and inhibition (accuracy on the Stroop task). Difficulty-adapted training also led to a reduction of number of head movements.

The findings from the preliminary study were replicated in a later double-blind, randomized study (Klingberg, et al., 2005) with a greater number of participants (n=44 completers). Although the design of the study resembled the design of the preliminary study it differed in certain aspects. Firstly, inclusion/exclusion criteria were different. The age span of the participating children was smaller (between 7 and 12); medication with psychoactive drugs was not accepted, etc. Secondly, certain adjustments had been made to the WMT program. It now consisted of a greater number of solely verbal and visuo-spatial working memory tasks. Thirdly, participants completed more tasks per day (90) and trained for a longer period of time each day (about 40 minutes). Mean number of training days was 25.2 (sd=2.2) in the treatment group and 26.6 (sd=2.6) in the comparison group. As in the earlier study, the control group trained with a non-difficulty-adjusted version of the same program as the training group used. Fourthly, new outcome measures (digit span and Conner’s Parent and Teacher Rating Scales) were added to the ones in the preliminary study. Fifthly, the later study contained three assessment points (pre-, post- and follow-up at three months) instead of two as in the earlier study. All significant results from the preliminary study were replicated except the observed decrease of head movements, which was non-significant in the later study. Furthermore the later study reported significant treatment effects on verbal working memory (digit span) and parent ratings of ADHD-typical behavior.

In the third published study (Holmes, Gathercole, Place, Dunning, Hilton, & Elliott, 2009) examining the effect of computerized WMT on children with ADHD, focus was turned to the issue of co-occurring medication. The age span of the participating children (n=25) was 8-11 years. All participating children were on either some form of methylphenidate or dexamphetamine medication. The procedure was as follows: first all children were assessed with neuropsychological tests while off medication (stopping at least 24 hours prior to assessment). Then the children were tested a second time on their usual medication. The training period started after the second assessment. The
computerized WMT program used was in this case essentially the same as used by Klingberg, et al. (2005). The children then trained for an average of 23.72 days over a period of between 6 and 10 weeks in school. The children completed 115 trials split across different tasks on each training occasion. All participants were then assessed post-training and at 6-month follow up. No control group existed in this study. Working memory was measured using the Automated Working Memory Assessment and performance and verbal IQ was measured using the Wechsler Abbreviated Scales of Intelligence. Significant treatment effects were found for both WMT and medication. The WMT led to improvements in all measured aspects of working memory. The results also indicated that the effects were maintained 6 months after training in three out of the four measured aspects of working memory (not verbal short-term memory). No effects were found on the measures of verbal and performance IQ.

A fourth unpublished study (Lucas, et al., 2008) was presented at the meeting of the American Psychiatric Association in May 2008. The study compared the effects of visuo-spatial and verbal WMT on a group of 46 children attending a summer camp for children with ADHD diagnosis. Results indicated that visuo-spatial WMT was superior to verbal WMT and that visuo-spatial WMT led to greater gains in working memory capacity and positive outcomes on behaviors as measured by a reward system used in the camp the children were attending.

Other studies concerned with computerized cognitive training have focused on non-ADHD patient groups and/or other forms of interventions. Holmes, Gathercole and Dunning (2009) investigated whether children with impairments in working memory would gain from the same sort of training as described by Klingberg et al. (2005). The children that trained with the difficulty-adapted working memory program showed improvements in all aspects of working memory and the ability to follow instructions at post-training assessment. All gains except in verbal short-term memory were maintained at a 6 months follow-up assessment. Moreover mathematical reasoning was improved at the follow-up.

The effects of computerized visuo-spatial WMT and inhibition training on a group of preschoolers have also been investigated (Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2008). WMT, but not inhibition training, showed effects on non-trained working memory tasks.

The same kind of WMT as described by Klingberg et al. (2005) has also been tried out on adults having suffered from a stroke (Westerberg, et al., 2007). The authors concluded that adaptive WMT led to improvements in working memory, attention and the subjective experience of cognitive functioning in everyday life in the examined population. Different kinds of computerized WMT programs have also been showed to improve performance on non-trained working memory tasks in old age (Li, Schmiedek, Huxhold, Smith, Lindenber, & Röcke, 2008; Buschkuehl, et al., 2008) and lead to a maintenance of effects on trained tasks for as longs as 18 months after training (Dahlin, Nyberg, Bäckman, & Stigsdotter Neely, 2008).

Jaeggi, Buschkuehl, Jonides, and Perrig (2008) have studied the effects of a different cognitive training program (based on the n-back task) that also contained the key feature
of adaptation of difficulty level based on prior performance. In their study healthy adults were subjected to a series of two sets of simultaneous stimuli: letters (auditory) and spatial locations marked on a computer screen (visuo-spatial). The participants were to decide if the combination of auditory and visuo-spatial stimuli had occurred on a specific position earlier in the series of parallel stimuli. Training with this alternative version of adapted WMT led to an improvement of working memory and fluid intelligence. An interesting finding was that the effects of training appeared to be dose specific; that is, more training resulted in bigger gains.

Another kind of computerized cognitive intervention that has commonalities to computerized WMT is computerized attentional training. Computerized training of attention has been associated with improvements in reading comprehension, passage copying and parents rating of inattention in children with ADHD (Shalev, Tsal, & Mevorach, 2007) and executive attention and intelligence in normal 4 and 6 year old children (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005). Both interventions contained some sort of adaptation of difficulty-level.

The Swedish Council on Technology Assessment in Health Care (SBU) recently published a review on scientific studies concerning computerized WMT and its effects on ADHD-symptoms. SBU stated that the number of scientific studies on computerized WMT was insufficient to draw conclusions on whether the intervention could decrease ADHD-symptoms or not. Several studies were excluded because they did not investigate the ADHD-population or did not meet criteria for scientific design according to the SBU-principles (SBU Alert, 2009).

Evidence in support of computerized WMT being capable of improving actual working memory capacity is growing. However the existence of transfer effects is more uncertain. As stated above there are indications that computerized WMT may lead to improvements in complex reasoning, inhibition (Klingberg, Forssberg, & Westerberg, 2002) and reductions of inattentive and hyperactive/impulsive behavior as rated by a parent (Klingberg, et al., 2005) in children with ADHD. These findings have so far not been replicated in other studies with participants fulfilling diagnostic criteria for ADHD.

It is a reasonable conclusion that more research needs to be done on the subject of computerized WMT in children with ADHD. The research referred to above lends many implications for which questions the present paper and future research should focus on. It is noteworthy that a substantial part of the studies uses slightly modified versions of the exact same computer program. It is difficult to conclude which aspects of the computerized training programs are necessary in order to improve working memory. One reasonable hypothesis is that the adaption of difficulty is a key feature of any computerized cognitive training program. One weakness of the entire body of research to date is the diversity in inclusion and exclusion criteria.
The aim of this study
The general aim of this study was to further explore the effects of WMT in children with ADHD. The participating children were randomized to 25 days of either computerized WMT (experiment condition) or computerized RT (control condition). The computerized RT was chosen as control condition in the present study since it was believed to be less tedious for the participants than the non-adaptive versions of the WMT that has sometimes been used as a control condition in previous studies (Klingberg, et al., 2005; Holmes, Gathercole, & Dunning, 2009; Klingberg, Forssberg, & Westerberg, 2002). The RT was also believed to be a more active control condition than the non-adaptive WMT, possibly affecting scholastic skills such as word - and letter decoding in a positive direction.

The ADHD population was of particular interest considering the previous promising results of WMT and the connections between ADHD and poor working memory. The high prevalence of ADHD, the difficulties associated with the disorder, the flaws of some of the frequent forms of treatment and the fact that ADHD does not automatically dissolve itself in adulthood constitute strong arguments for developing additional treatments for the disorder.

Improvements of working memory could be expected to have positive effects on episodic memory and verbal learning, not by improving the long-term store per se, but by optimizing initial processing and encoding of information. The connection between working memory capacity and mathematical performance and reading also makes it probable that WMT could lead to improvements in these areas. If working memory can be considered a fundament for more complex, goal-directed behavior, as suggested by Barkley (1997), then WMT should also have an effect on behavioral symptoms.

The tests chosen as outcome measures were meant to assess the above mentioned problems associated with ADHD, namely working memory deficits, poor scholastic skills and behavioral symptoms. Tests were assessed in a preliminary evaluation study prior to the recruitment of participants in the actual study. Selection of participants was made striving for a high ecological validity. Therefore common comorbid disorders or stimulant medication were not considered as exclusion criteria.

Three hypotheses were tested in the present study:
1. WMT improves working memory capacity in children with ADHD.
2. WMT improves scholastic skills in children with ADHD.
3. WMT decreases diagnostic symptoms in children with ADHD.
Method

The present study was approved by the regional ethical review board in Stockholm. All participating children were informed about the voluntary nature of the study, that all research data would be treated with confidentiality and that they were free to terminate participation at any given time. Written consent was obtained from all participating parents and children.

Participants

38 potential participants were screened on telephone and 32 met the inclusion/exclusion criteria (Table 1, Figure 2). The participating families were instructed to bring summaries of neuropsychological assessments, stating the children’s DSM-IV diagnosis. All reported diagnoses were based on these summaries. SNAP parent ratings (see Outcome measures) were used to assess current symptomatology. Since most children were on medication, the SNAP ratings were not thought to reflect the actual degree of difficulties concerning inattention and/or hyperactivity/impulsivity and the SNAP scores were thus not used in the inclusion/exclusion procedure. Despite medication, the majority of participants (n=25) were rated above the cut-off value for the combined subscale (≥1.67). Of the remaining 7 participants, 3 were rated above the cut-off value for the inattention subscale (≥1.78), 2 for the hyperactive/impulsive subscale (≥1.44) and 2 were not rated above any of the cut-off values. The cut-off values were recently validated in a study of the psychometric properties and normative values of the SNAP scale (Bussing, et al., 2008).

The inclusion criteria were: 1. ADHD-diagnosis (including all three different subtypes and Hyperkinetic disorder). 2. Year of birth 1999-2003. 3. Access to PC computer at home. The exclusion criteria were, 1. Previous structured computerized cognitive training for more than five days. 2. Mental retardation diagnosis. 3. Autism diagnosis (Asperger syndrome and pervasive developmental disorder were not considered as exclusion criteria). 4. Change of dose, or starting use of Methylphenidate, Atomoxetine and/or Prometazin five weeks before the trial and until the end of the study.

Table 1. Participant characteristics

<table>
<thead>
<tr>
<th></th>
<th>WMT</th>
<th>RT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td>7 (6)</td>
<td>11 (3)</td>
<td>18 (9)</td>
</tr>
<tr>
<td>Girls</td>
<td>1 (1)</td>
<td>2 (1)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>ADHD combined</td>
<td>8 (7)</td>
<td>11 (3)</td>
<td>19 (10)</td>
</tr>
<tr>
<td>ADHD inattentive</td>
<td>0 (0)</td>
<td>1 (1)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Hyperkinetic disorder unspecified</td>
<td>0 (0)</td>
<td>1 (0)</td>
<td>1 (0)</td>
</tr>
<tr>
<td>Aspergers syndrome</td>
<td>1 (0)</td>
<td>1 (0)</td>
<td>2 (0)</td>
</tr>
<tr>
<td>Pervasive developmental disorder NOS</td>
<td>0 (1)</td>
<td>1 (0)</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Developmental coordination disorder</td>
<td>1 (0)</td>
<td>2 (0)</td>
<td>3 (0)</td>
</tr>
<tr>
<td>Tourette’s syndrome</td>
<td>2 (0)</td>
<td>0 (0)</td>
<td>2 (0)</td>
</tr>
<tr>
<td>Phonological &amp; grammatical LI</td>
<td>1 (0)</td>
<td>0 (0)</td>
<td>1 (0)</td>
</tr>
<tr>
<td>Methylphenidate</td>
<td>6 (4)</td>
<td>7 (2)</td>
<td>13 (6)</td>
</tr>
<tr>
<td>Atomoxetine</td>
<td>1 (0)</td>
<td>2 (2)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Mean age, years</td>
<td>8.25 (8.75)</td>
<td>8.75 (9)</td>
<td>8.50 (8.83)</td>
</tr>
<tr>
<td>Mean time passed since diagnosis, years</td>
<td>1.51 (2.31)</td>
<td>1.87 (0.58)</td>
<td>1.73 (1.68)</td>
</tr>
<tr>
<td>Computer use, hours per week (m)</td>
<td>5.94 (7.93)</td>
<td>7.81 (11.75)</td>
<td>7.10 (9.32)</td>
</tr>
</tbody>
</table>

*aData for completers (to the left) and for drop-outs (within parentheses).
Recruitment
Children were recruited through an ADHD-specialized information center, a national support group for neuropsychological disorders, elementary schools and children’s psychiatric clinics. Information regarding the study was sent via mail to all parents that were enlisted at the ADHD-specialized information center and who had children in the target age group. The letter contained a short description of the study and contact information for getting more information concerning the project. The same letter was also placed in waiting-rooms at the children’s psychiatry clinics and was sent from schools to potential participants parents. Finally, information was posted on the webpage of the above mentioned national support group for parents of children with neuropsychological disorders. The parents were instructed to contact the instigators of the study for further information.

Procedure
Each parent who contacted the instigators of the study and whose child was considered to be a possible participant was screened for inclusion/exclusion criteria in a structured telephone interview. Parents whose children were included in the study were asked to keep dosing of their children’s medication stable until post testing and not to start any new medication until after post assessment. The included participants were randomized to a WMT group (n=15) and a RT group (n=17). Each child was assessed pre-training (m=5.5 days pre-training, sd=2.8) and parents participated in one hour of education in adjunction to the first assessment. A 5 to 8 week long period of computerized training started within one week after the first assessment. Post-training assessment followed as soon as possible after completed training (m=5.8 days after training, sd=4.0). See Figure 2 for a schematic overview of the design of the study.
Each child was randomized to be assessed by one of two assessors. The assessors were graduate clinical psychology students. Pre- and post-testing for each child was conducted by the same assessor in order to control for effects of the relation between administrator and child. The assessors had previously trained in administering all the outcome measures on a total of ten occasions each with five undiagnosed, nine year old children, at an elementary school.

One hour was dedicated for each assessment. While the children were assessed the parents filled out two rating scales in a waiting-room. At the first assessment parents also filled out a form with background data. Children were scheduled to test at the same time of the day at all assessments to control for medication effects and circadian rhythm. The tests were performed in a quiet room with only the child and the assessor present. At a specified point in the assessment a small break was planned. During the break the assessor and the participating child stayed in the testing room. The assessments lasted in average 33 minutes (sd=4.1) At the second assessment parents and children were instructed not to give information to the assessors regarding what type of training program that was used to ensure that the assessors were blind to which form of training the child participated in. After each assessment parents were given a teacher rating scale for the child’s primary teacher to fill out in school. The rating was then brought to the next assessment.

Evaluation of outcome measures
Preliminary evaluation of the outcome measures was conducted at an elementary school with a group of 9-year old children (n=10). All children attended the same class and were selected to participate by their primary teacher. The teacher was instructed to only select children without a documented neuropsychiatric diagnosis. The children were assessed at their school on two separate occasions with 5 weeks in-between. All parallel versions of the tests were used in the evaluation and no child was assessed with the same version on pre- and post-assessment. The evaluation served two purposes: Firstly to train the assessors on administration procedures and secondly to investigate test-retest effects and statistical characteristics of the outcome measures. Based upon this evaluation the primary outcome measure on the word span task was altered from maximum correct words in a series to total number of correct trials since the childrens’ results on maximum correct words in a series were too homogenous. Block-tapping was replaced with a computerized span-board task, primarily to standardize test administration. The other tests used in the evaluation were the same as in the final study. No parent or teacher ratings were administered. Paired samples t-tests revealed that there were no significant differences between mean results at pre- and post-assessment on any of the outcome measures, except Letter-chains and Word-chains (see Table 2). The children participating in the evaluation of the outcome measures completed significantly more chains on Letter- and Word-chains at the second assessment than at the first.
Table 2. Mean values, standard deviations and t-values on paired samples t-test of the first and second assessment in the outcome measures evaluation.

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Pre m</th>
<th>Pre sd</th>
<th>Post m</th>
<th>Post sd</th>
<th>T-test t&lt;sub&gt;o&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVLT words reproduced</td>
<td>44.3</td>
<td>9.72</td>
<td>44.8</td>
<td>7.60</td>
<td>-0.885</td>
</tr>
<tr>
<td>AVLT delayed recall percent recollected</td>
<td>86.2</td>
<td>24.7</td>
<td>78.2</td>
<td>20.3</td>
<td>1.02</td>
</tr>
<tr>
<td>Word span max. words recollected</td>
<td>3.90</td>
<td>0.738</td>
<td>4.00</td>
<td>0.667</td>
<td>-0.361</td>
</tr>
<tr>
<td>Word span correct trials</td>
<td>4.90</td>
<td>1.45</td>
<td>5.00</td>
<td>1.05</td>
<td>0.176</td>
</tr>
<tr>
<td>Arithmetic highest level achieved</td>
<td>11.5</td>
<td>2.88</td>
<td>12.1</td>
<td>1.73</td>
<td>-0.627</td>
</tr>
<tr>
<td>Letter-chains correct chains</td>
<td>30.5</td>
<td>4.99</td>
<td>36.1</td>
<td>4.12</td>
<td>-5.59*</td>
</tr>
<tr>
<td>Word-chains correct chains</td>
<td>25.2</td>
<td>6.48</td>
<td>29.8</td>
<td>6.58</td>
<td>-4.01*</td>
</tr>
</tbody>
</table>

*p<0.01

Outcome measures
The order in which the tests were administered was the same for pre- and post-intervention assessment. The intention for a set order was to control for such factors as tiring during testing and getting approximately the same amount of time between the verbal learning test (AVLT) and the delayed recall. The administration order of the tests was: 1. AVLT. 2. Span-board. 3. Arithmetic. 4. Word span. 5. Letter-chains. 6. Word-chains. 7. AVLT delayed recall. Parallel, similar but on certain key-features different, versions were created for some of the tests – (a), (b), (c), and (d) – in order to minimize test-retest effects. Randomization determined which of the parallel versions the participants were to complete at each assessment point, so that the effect of possible variations in the degrees of difficulty of the parallel versions of tests (a-d) was reduced.

Working memory
(a) A computerized span-board task was used to measure visuo-spatial short term memory. Red dots were presented in sequences in a four-by-four grid. After each sequence was finished the child was to reproduce the sequence by using the computer mouse and clicking in the same squares and in the same order that the red dots had appeared. Initially two different sequences consisting of two dots were presented as a trial round. Then two sequences consisting of two dots, two sequences with three dots and so forth were presented. The test was discontinued when the child failed to reproduce two sequences on the same level of difficulty. The interstimulus time for this test was set for 750 ms and the red dots appeared on the screen for 2250 ms. The outcome measure for the span-board task was total number of correct clicks.

(b) A word span task was administered in order to measure verbal working memory. This test is a variant of the Digit Span subtest from WISC-III (Wechsler, 1991) in which the test person is to repeat a serious of nouns instead of numbers. The test was constructed as described in Thorell and Wåhlstedt (2006), but only using the backward condition. As with the AVLT-adaptation described above three separate versions of the test were constructed. Each wordlist used the same set of monosyllabic Swedish nouns, but the words were randomized in different orders for each version. The nouns were recorded by the same person as in the AVLT in wav-files, which were then played for the child through headphones with 1000 ms interstimulus time. At first the child was to repeat 2 words in a reversed order 2 times, then 3 three words 2 times and so on. The test was discontinued when the child failed both trials on the same difficulty level. Number of correctly repeated trials was the outcome measure for the word span task.
Scholastic skills

(c) Verbal learning was assessed using the Auditory Verbal Learning Test (AVLT; Lezak, 1995). Three parallel wordlists with 15 frequently occurring Swedish nouns in each was created. Each child was tested with a different word list on each test occasion. All 45 words were recorded in wav-files by a person who otherwise didn't participate in the study. The child was instructed to put on headphones before the test started. The words were then played with 1000 ms interstimulus time. The administration of the test differed somewhat from the standard administration of the AVLT. The word span task (b) was considered to be a sufficient interference and therefore only one wordlist was used in the AVLT per assessment. The outcome measure of the AVLT was total number of correctly reproduced words over the 5 trials.

(d) An arithmetic test was constructed according to the same pattern as the subtest Arithmetic from WISC-III (Wechsler, 1991). The test was used in order to assess mathematical reasoning and applied verbal working memory (remembering the question asked while simultaneously working on the solution). All questions were administered verbally without any visual support available. Three parallel versions of this test were created. Numbers were slightly changed between each parallel version of every question, with caution in order not to change the difficulty-level. Nouns were also replaced so that parallel questions seemed more different. An example of two parallel versions of a question was: “How many are 4 pencils and 5 pencils together?”, and “How many are 3 rubbers and 4 rubbers together?” The questions were ordered in ascending difficulty level. The outcome measure of the arithmetic test was the number of the most difficult question solved.

(e) Letter-chains (Bokstävskedjor) from Reading-chains (Läskedjor; Jacobson, 2001) was intended to measure letter decoding speed. In this task the child was to visually search chains of letters (e.g. AEKKFJEEN) for letter repetitions (e.g. KK) and to draw a line between the letters in every occurring letter pair. The child was to finish as many letter-chains as possible during two minutes. The measurement in this test was the total number of chains with correctly drawn lines.

(f) Word-chains (Ordkedjor) from Reading-chains (Jacobson, 2001) was used to assess word-decoding ability. In this task the children were instructed to scan chains consisting of 3 words of 2-4 letters each (e.g. catballmum) and then draw lines where one word ended and the next started. The child was to finish as many word-chains as possible in 2 minutes time and the total number of chains with correctly drawn lines were then counted and used as the outcome measure of the test.

(g) Storage and retrieval of episodic memories was measured by dividing the maximum number of recollected words in a single trial with the number of correctly recollected words on the delayed recall part of the AVLT (Lezak, 1995).

Diagnostic symptoms

(h) SNAP Parents and Teachers Rating Scale (Swanson, et al., 2006) was used in order to assess ADHD-symptoms as described in the DSM-IV (American Psychiatric Association, 1994). The Swedish version of SNAP was obtained from the national registry of treatment evaluation for ADHD (Swanson, n.d.). Parents and teachers were
to rate 18 statements concerning the child on a 4-point Likert scale, from 0 (“Not at all”) to 3 (“Very much”). High ratings reflect presence of hyperactivity/impulsivity and inattention. The outcome measures were the total sum of the nine items concerning inattention and the total sum of the nine items concerning hyperactivity/impulsivity.

(i) Leiter-R Parents Rating Scale (Roid & Miller, 2001) was used to assess other aspects of behavioral symptoms relating to ADHD. The scale consisted of 8 subscales constituting 2 composite scales with a total of 51 items. Each item consisted of descriptions of 2 opposing behaviors and the parents were to choose if either of the behaviors were more typical for their child or if both behaviors were common. Low scores reflect difficulties. The first composite scale (Cognitive/Social) consisted of the 4 subscales Attention, Level of Activity, Impulsivity and Social Ability. The second composite scale Emotions/Regulation consisted of the 4 subscales Adaptation, Mood and Self-confidence, Energy and Emotions and Sensibility and Self-regulation. The outcome measures were the raw score on the Cognitive/Social composite scale and the attention subscale.

Trained tasks

(j) The parents of the participants in the RT group were instructed to copy the number of attempts and correct responses from the computer software for each trial executed during training. The total number of correct responses was then divided by the total number of attempts in the first and last 5 days of training respectively. In order to assess changes in the performance on the trained tasks in the RT condition the resulting ratios were compared.

(k) Maximum number of recollected items forwards and backwards for each trained task was registered in the computer software. Since the number of tasks backwards differed somewhat between participants, no calculations were made on the backward tasks. Mean values on all forward tasks in the first and last 5 days of training were computed and used to measure effects on the trained tasks in the WMT condition.

Interventions

At least one parent of each participating child was required to attend one hour of education where instructions for using the different computer programs were presented. The educational meeting also included instructions for how to administer rewards for completed training weeks, how to coach the child and give feedback during training, and how to register training on a training schedule. The education was conducted by a certified psychologist, who is one of the designers of the programs. Separate but similar educations were given depending on which group the parent’s child was randomized to, for two reasons. The first to keep the parents blind regarding what intervention the other group received and the second that information specific for each program was given.

Both training interventions were performed at home with the child and parent training together. The training was scheduled to last 5 days a week, 5 consecutive weeks. At least 20 days of training during a period of at the most 8 weeks was a requirement for final inclusion in the study. Each parent was phoned 2 times during training by the certified psychologist who was in charge of the educations, to follow up on the process of training and to answer questions if the parents had any.
Both training interventions included a reward-system external to the training programs. Each parent was to decide a suitable reward for each completed week of training to increase compliance and motivation in the children.

**Working memory training**

The children in the WMT group used Memory Games Senior (Läramera Program AB; Leripa AB; Kognitiva Kompaniet AB, 2008), a computer program designed for training working memory according to the same basic principles as other working memory programs that have shown positive effects on working memory (Backman & Truedsson, 2008; Holmes, Gathercole, & Dunning, 2009; Klingberg, et al., 2005; Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2008).

The program contained 9 tasks that were to be completed on each day of training. The number of trials per task was 10. Average training time was 44.3 minutes (sd=10.6) per training day. The general objective of the training tasks was to memorize and reproduce the order of a number of presented items. The tasks differed in lay-out, type of items presented, if the items were presented visually or verbally and if the items were visible during presentation. The item presentation order was randomized in all tasks. The difficulty of all tasks was continuously adapted depending on the child’s performance. After a set number of correct responses the items to be remembered increased. The levels were also adjusted down if a set number of incorrect responses were made. The purpose of this adaptation was to constantly maximize the load on working memory. If the child reached a certain level of difficulty the child was instead required to click on the items in reverse order. The reverse condition applied to about half of the exercises.

In tasks 1, 2 and 3 the items presented were visual on screen (fish blowing bubbles, lamps blinking and Egyptian symbols being marked). Each task had only 1 sound corresponding to all items in the task. Sequence memorization based on sound was therefore not possible. In tasks 4 and 5 items (letters and colors) were presented verbally while not visible on screen. Immediately after presentation, visual items equivalent to the verbally presented stimuli and distractors appeared on the screen. The sixth task consisted of a piano playing random melodies. The keys moved as the melodies were played and the keys had letters representing the tones written on them. Task 7 contained verbally presented letters and digits in alternate order (see Figure 3). The object of the task was to recall the letters first in correct order, and then the digits in correct order, requiring the child to sort the presented stimuli into categories before recalling them, while at the same time keeping track of the positions of each item. Tasks 8 (see Figure 3) and 9 contained visual items that each had a unique sound (i.e. a lion roaring, a pianist playing the piano). In the eighth task the items had fixed positions on screen and in the ninth exercise localization were randomized.
Feedback was included in the program in several forms. After each correct response an encouraging voice gave the child a positive comment, and after an incorrect response the child was encouraged to keep trying. Each training day after the nine tasks were completed the child was allowed to play a computer game included in the program for approximately 5-10 minutes as a reward. Each correct click in the training gave one point and the points were then converted to lives used in the reward game. The program also kept track of the best results in both the reward game and on each task and biggest improvements on a single task for each day of training to further facilitate motivation.

A log-book was used which contained a schedule for the external rewards, some information about WMT and an overview for each week of training where the best results and improvements for each day of training were to be entered. Also a training schedule was used where the time of training for each day, and on what dates the training took place were to be entered. Detailed results were also stored in a file in the computer program. To make sure all necessary information was provided a checklist was used guiding through all steps that were required to be taken on each day of training.

Reading training
The children in the RT group used Reading World (Läramera Program AB; Leripa AB, 2003), a program designed for language and literacy development. The program consisted of 21 exercises divided into three categories: Practice Whole Words, Practice Letters and Build Words.

Practice Whole Words consisted of 6 different exercises, in which the child was to practice for example sorting words in alphabetical order (see Figure 4), matching a word to the right picture amongst a number of pictures, and matching a picture to the correct word.

Practice Letters contained 6 different exercises that allowed the child to practice for example selecting a picture amongst three where the first letter of the selected picture matches a letter presented above the pictures, finding the correct letter on the keyboard when a letter is presented on the screen (see Figure 4) and sorting letters in a scrambled word to form the correct word.
Build Words consisted of 9 different exercises that required the child to practice for example puzzling together words from separated and scrambled syllables, building a word letter by letter and figuring out what letters are included in a word.

Figure 4. Screen shots from a task in the Practice Letters category (to the left) and a task in the Practice Whole Words category (to the right).

Wordlists of about 400 different words were included in the program and were used in all of the exercises. Feedback was included in Reading World in slightly different forms. Correct responding was followed by an encouraging melody and incorrect responding resulted in a toot-like sound.

The participants randomized to the RT were instructed to undertake one exercise from the Practice Whole Words category, one exercise from the Practice Letters category and two exercises from the Build Words category for about 5 minutes per exercise and training day. Average training time was 25.3 minutes (sd=6.2) per training day.

To keep track of the training, a schedule was used where time for training, what dates training took place and results on the different exercises were to be entered day by day. A separate form was used to keep track of the external rewards. To make sure all necessary information was provided a checklist was used guiding through all steps that were required to be taken on each day of training.

The control condition in the present study was meant to resemble the WMT condition in several aspects without specifically training working memory. Both computer programs used were made by the same illustrator, pedagogue and programmer. They were commercially available and used in Swedish elementary schools at the time of the study. The programs were not adapted from the original commercially available versions for the purpose of this study. RT was also chosen because reading disabilities is a commonly occurring problem in the ADHD population. Therefore it was assumed that RT would have high face validity, and that parents and children would be motivated to complete training. The intent was for both interventions to be perceived as active treatments for all participants.
Statistical analysis
One-tailed paired samples t-tests were performed for all tests and rating scales comparing pre- and post-intervention scores in the WMT and RT group separately. Delta-values were calculated by subtracting the pre-intervention score from the post-intervention score. One-tailed independent samples t-tests were performed comparing the delta-values between the WMT and RT condition. The choice of one-tailed t-tests was made in accordance with the hypothesis of the study and based upon results from previous studies. There is nothing that indicates that WMT could have adverse effects on any of the outcome measures.

In order to investigate differences on the trained task, one-tailed paired samples t-tests were performed comparing results on the first and last 5 days of training. Results on the first 5 days of training were then subtracted from results on the last 5 days of training in order to obtain delta-values. Results on the trained tasks were ranked based upon scores on the first 5 days of training and divided into top half and bottom half in each condition. One-tailed independent samples t-tests were performed comparing the delta values of the top and bottom half in each condition separately. Results on the trained tasks were investigated to assess if possible effects on tests and rating scales were linked to improvements on the trained tasks.

Z-scores were calculated on all tests and rating scales with tendencies or significant differences within or between groups. Raw mean of pre-intervention assessments were used as an estimate of the populations mean.

Results
Some of the statistical analyses were calculated with less than the maximum observed values. The data was lost due to different reasons. A few (n=4) of the participators could not read or refused to execute Word-chains and values for 1 participant on the span-board task was lost due to computer malfunction. A number of parents (n=2) failed to fill out some of the items in Leiter-R and 1 parent failed to fill out the SNAP and Leiter-R rating scales. A number of teachers (n=8) did not fill out or return the SNAP rating scale.

To investigate possible dissimilarities between the WMT and RT groups, independent samples t-tests with all outcome measures, average computer use per week, age, number of days of training and total number of days between first and last day of training as dependent variables and type of intervention as independent variable was performed. Fisher’s exact test was performed for type of ADHD-diagnosis, all co-morbid disorders, medication, gender and special support in school. No statistically significant differences were found on any of the background variables or pre-intervention assessment scores.

Raw data for all tests and rating scales and t-scores for all t-tests are presented in Table 3. Z-scores are presented in Figures 6-8. Table 4 contains data for the trained tasks.
Table 3. Outcome measures at pre- and post-intervention assessment.

<table>
<thead>
<tr>
<th>Measure</th>
<th>WMT</th>
<th>RT</th>
<th>Pre</th>
<th>Post</th>
<th>Pre-post t-test&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Delta t-test&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVLT words reproduced</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=33.8 (4.98)</td>
<td>37.3 (11.1)</td>
<td>-0.911</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=35.8 (9.10)</td>
<td>m=36.8 (11.2)</td>
<td>-0.632</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Span-board correct clicks</td>
<td></td>
<td></td>
<td>n=7</td>
<td>m=22.9 (12.0)</td>
<td>32.9 (13.4)</td>
<td>-1.81&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=17.8 (13.1)</td>
<td>m=18.7 (16.0)</td>
<td>-0.553</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Span-board reaction time&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>n=7</td>
<td>m=535.8 (97.4)</td>
<td>537.0 (79.0)</td>
<td>-0.047</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=520.0 (105.0)</td>
<td>m=520.0 (92.3)</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic highest level achieved</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=7.63 (4.21)</td>
<td>10.0 (3.38)</td>
<td>-2.30&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=8.53 (4.84)</td>
<td>m=8.53 (3.73)</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word span task correct trials</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=3.13 (3.54)</td>
<td>4.00 (1.20)</td>
<td>-2.20&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=3.08 (0.954)</td>
<td>m=3.08 (1.50)</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter-chains correct chains</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=25.3 (9.51)</td>
<td>30.6 (13.2)</td>
<td>-2.47&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=25.5 (10.6)</td>
<td>m=31.7 (10.7)</td>
<td>5.87&lt;sup&gt;**&lt;/sup&gt;</td>
<td>2.36&lt;sup&gt;**&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Word-chains correct chains</td>
<td></td>
<td></td>
<td>n=7</td>
<td>m=12.9 (7.82)</td>
<td>17.0 (9.68)</td>
<td>-4.10&lt;sup&gt;**&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=10</td>
<td>m=15.1 (10.6)</td>
<td>m=18.3 (12.6)</td>
<td>-2.50&lt;sup&gt;&lt;/sup&gt;</td>
<td>1.90&lt;sup&gt;&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>AVLT delayed recall percent recollected</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=75.4 (20.6)</td>
<td>66.7 (11.8)</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>n=13</td>
<td>m=76.3 (22.2)</td>
<td>m=73.7 (25.0)</td>
<td>0.376</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNAP parent hyperactivity/impulsivity</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=16.1 (2.47)</td>
<td>12.5 (3.74)</td>
<td>2.83&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=12</td>
<td>m=16.4 (5.33)</td>
<td>m=14.9 (6.30)</td>
<td>0.977</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNAP parent inattention</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=14.5 (5.48)</td>
<td>11.7 (3.58)</td>
<td>2.71&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=12</td>
<td>m=18.1 (4.14)</td>
<td>m=15.3 (3.45)</td>
<td>1.61&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.34&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>SNAP teacher hyperactivity/impulsivity</td>
<td></td>
<td></td>
<td>n=5</td>
<td>m=11.2 (8.56)</td>
<td>13.4 (7.47)</td>
<td>-1.28</td>
</tr>
<tr>
<td></td>
<td>n=8</td>
<td>m=12.1 (7.62)</td>
<td>m=13.9 (9.78)</td>
<td>-1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNAP teacher inattention</td>
<td></td>
<td></td>
<td>n=5</td>
<td>m=14.6 (3.91)</td>
<td>14.2 (4.92)</td>
<td>.266</td>
</tr>
<tr>
<td></td>
<td>n=8</td>
<td>m=15.6 (6.90)</td>
<td>m=17.0 (8.18)</td>
<td>-.743</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEITER-R attention</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=17.6 (3.74)</td>
<td>18.9 (3.18)</td>
<td>-1.67&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=12</td>
<td>m=14.7 (3.37)</td>
<td>m=16.3 (3.14)</td>
<td>-1.66&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.17&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>LEITER-R cognitive/social</td>
<td></td>
<td></td>
<td>n=8</td>
<td>m=48.8 (7.32)</td>
<td>51.9 (4.58)</td>
<td>-1.82&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>n=10</td>
<td>m=44.4 (7.28)</td>
<td>m=46.8 (6.30)</td>
<td>-1.08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Paired samples t-tests comparing means at pre- and post-intervention assessment (one-tailed).
<sup>b</sup>Independent samples t-tests comparing the pre-post change (delta) in the WMT condition with the pre-post change in the RT condition (one-tailed).
<sup>c</sup>Milliseconds.
<sup>d</sup>p<.1
<sup>*</sup>p<.05
<sup>**</sup>p<.01
Hypothesis 1: WMT improves working memory capacity in children with ADHD

The performance of the WMT group increased significantly on the word span task from pre- to post-intervention assessment (see Table 3). The gain of the WMT group was significantly greater than the change in the RT group. The effect size of the difference between the groups on the word span task was calculated using Cohen’s delta and was \(d=0.89\). There were no significant improvements in either group on the span-board task, but there was a tendency in favor of WMT, both when comparing number of clicks at pre- and post-intervention assessment and when comparing the improvement of the groups. The effect size of the tendency towards improvement was \(d=0.80\). The results indicate that WMT leads to greater improvements of verbal working memory performance than RT, in part supporting hypothesis 1.

![Figure 6. Mean values on tests measuring working memory capacity at pre-training and post-training.](image)

Hypothesis 2: WMT improves scholastic skills in children with ADHD

When comparing mean scores on the arithmetic test at pre- and post-intervention assessment, the WMT group significantly improved, while the RT group scored equally at both assessment points (see Table 3). An analysis of the change of the groups’ performances revealed that the WMT group improved significantly more than the RT group (\(d=0.89\)). There were no other statistically significant differences between the delta values of the groups on any of the other tests measuring different aspects of scholastic skills. However, both groups significantly improved on Letter- and Word-chains. In conclusion, WMT lead to significantly greater improvements on mathematical reasoning than RT, thereby lending support to part of hypothesis 2.

![Figure 7. Mean values on tests measuring different aspects of scholastic skills at pre-training and post-training.](image)
Hypothesis 3: WMT decreases diagnostic symptoms in children with ADHD

No significant differences were found between the delta values of the groups on any of the parent or teacher rating scales (see Table 3). The participants in the WMT group were rated by their parents to have a significantly decreased prevalence of behavioral symptoms of both inattention and hyperactivity/impulsivity as measured by SNAP after the intervention. There were tendencies of symptom reduction in the WMT group on the Leiter-R inattention and cognitive/social scales as well as tendencies of both measures of inattention as rated by parents in the RT group. In summary, hypothesis 3 was not supported in the present study. There was no significant difference between the RT and WMT groups regarding decrease of ADHD-symptoms.

![Figure 8](image.png)

Figure 8. Mean values on parent ratings of behavioral symptoms at pre-training and post-training assessment points. In the SNAP rating, lower scores indicates decrease of symptoms, and in the Leiter rating, higher scores indicates decrease of symptoms.

Effects on the trained tasks

The WMT group performed significantly better on the trained tasks during the last 5 days of training compared to the first 5 days of training (see Table 4). There were no significant differences on the mean improvement of results on the trained tasks when comparing the top and bottom half of the WMT group. However, the bottom half, but not the top half improved significantly. Since no differences between top and bottom half in the WMT group were found, no further statistical analysis on the subgroups were performed. In contrast the RT group did not generally improve on the trained tasks. The bottom half significantly improved performance on the last 5 days of training compared to the first 5 days of training and the improvement was significantly greater than the top half. Although not statistically significant, the top half somewhat lowered performance on the trained tasks.
Table 4. Results on trained tasks at first and last 5 days of training.

<table>
<thead>
<tr>
<th>Working memory training</th>
<th>First 5 days</th>
<th>Last 5 days</th>
<th>Pre-post t-test⁹</th>
<th>Delta t-test⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>m (sd)</td>
<td>m (sd)</td>
<td>t-test</td>
</tr>
<tr>
<td>Maximum items</td>
<td>7</td>
<td>4.87 (.601)</td>
<td>5.71 (.758)</td>
<td>-4.91**</td>
</tr>
<tr>
<td>Top half</td>
<td>3</td>
<td>5.42 (.399)</td>
<td>6.12 (.542)</td>
<td>-2.15§</td>
</tr>
<tr>
<td>Bottom half</td>
<td>3</td>
<td>4.33 (.218)</td>
<td>5.18 (.319)</td>
<td>-3.36</td>
</tr>
<tr>
<td>Reading training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent correct trials</td>
<td>12</td>
<td>74.2 (2.81)</td>
<td>77.3 (13.0)</td>
<td>-.799</td>
</tr>
<tr>
<td>Top half</td>
<td>6</td>
<td>81.9 (4.43)</td>
<td>77.8 (12.8)</td>
<td>.863</td>
</tr>
<tr>
<td>Bottom half</td>
<td>6</td>
<td>66.6 (6.94)</td>
<td>76.9 (14.5)</td>
<td>-2.23*</td>
</tr>
</tbody>
</table>

⁹Paired samples t-tests comparing means at pre- and post-intervention assessment (one-tailed).
§Independent samples t-tests comparing the pre-post change (delta) in the bottom half with the top half in the WMT condition and in the RT condition (one-tailed).
**p<.1
§p<.05
†p<.01

Drop-out analysis

For the participants who withdrew before post-training assessment, a drop-out analysis was conducted. Independent samples t-tests were performed for all outcome measures comparing mean values of completing participants and participants who withdrew showing no statistical differences on any of the outcome measures at pre-intervention assessment.

Discussion

This study explored the effects of adaptive WMT on working memory, scholastic skills, and ADHD-symptomology in children diagnosed with ADHD. An intensive training period, consisting of at least 20 training occasions spread across 5-8 weeks, lead to improvements in verbal working memory and mathematical reasoning.

The most important new finding of this study was the improvements concerning mathematical reasoning. It should be acknowledged that there might be superior measures of arithmetic skills, such as grades or results on national tests but neither was available in the present study. The improvements in mathematical reasoning are of special interest since mathematical disabilities are frequently seen in children with ADHD (Faraone, et al., 1993). Solving verbally presented arithmetic problems involves keeping information in the phonological loop, retrieving arithmetic rules from long term memory and using the information to solve the problem at hand. It is thus reasonable that gains in verbal working memory can lead to improvements in mathematical reasoning.

The improvements in verbal working memory apparent in the present study is in line with findings in earlier studies of the effects of WMT on ADHD-children (Klingberg, et al., 2005; Holmes, Gathercole, Place, Dunning, Hilton, & Elliott, 2009). Increasing verbal working memory capacity would be helpful for many children with ADHD, since they often display difficulties with following instructions, developing reading skills and other areas involving verbal working memory.
Visuo-spatial working memory did not significantly improve, although the results indicated a tendency for the WMT group to have a larger effect. There are at least three possible explanations for the lack of significant gains in visuo-spatial working memory. Firstly, the span-board task only required forward repetition. It could be argued that forward repetition only includes storing of information and not processing and thus only measures one of the key aspects of working memory. Secondly, although the effect size was large on the span-board task, standard deviations in the groups were substantially unequal and thus no significant effects were seen. When calculating effect size based upon the higher standard deviation of the WMT group, 13 participants in each group would have been needed to ensure statistically significant effects, given that the results would continue in the same direction. A third explanation is that there are no significant differences regarding the gains of the two groups. This third explanation seems unlikely based on the fact that the span-board task is similar to the trained tasks in the WMT group and that the WMT group significantly improved performance on the trained tasks.

It is difficult to say whether improvement of performance on working memory tasks following WMT reflects an actual improvement of working memory capacity or the acquisition of new strategies. The word span task is slightly similar to the task with verbally presented digits in the WMT. It would be hard to reject the notion that effects are merely due to development of new strategies if the improvements were limited to the working memory outcome measures. The fact that the children in the WMT group also improved on the arithmetic test make it less plausible to conclude that strategy development is the only explanation of the improvements. Since improvements were found on tasks similar to those in the WMT intervention and tasks less similar, it is more likely that the improvements are due to actual gains of working memory capacity. Nonetheless even if the improvements could be fully explained by strategy development, these strategies seem applicable in diverse situations and the argument of strategy-development would hardly be in disfavor of WMT. One possible explanation to why training of a specific cognitive function such as working memory can lead to changes in other functions, suggested by Olesen, Westerberg and Klingberg (2004), is that the cortex that is affected by the computerized WMT can be considered multimodal, and thus relating to more than just one cognitive function.

The significant decrease of parent-rated inattentive symptoms of ADHD found in an earlier study (Klingberg, et al., 2005), were not replicated in the present study. The tendency was for both groups to decrease symptoms of inattention, and the WMT group was rated to have significantly less symptoms of hyperactivity and impulsivity at post-intervention assessment than pre-intervention assessment as rated by parents. These results allow different interpretations. It is possible that both the RT and WMT conditions affect ADHD-symptomatology in a positive direction. Also, the parents of all participating children have gained the experience of seeing their children completing a repetitive task on a number of occasions. This could explain the fact that the parents tended to rate their children as more attentive.

It’s noteworthy that the arithmetic test is, as most neuropsychological tests, not a measure of one specific cognitive function. Working memory capacity is essential in solving both sorts of tasks. Provided that the child masters the basic rules of arithmetic, the arithmetic test can be considered essentially a test of applied working memory. The
child is required to keep the question in the phonological loop while working on solving the problem. The fact that the demands on working memory is much greater concerning the arithmetic test compared to Word-chains, Letter-chains, or AVLT is one possible explanation to the lack of effects on the later tests. Another possible explanation of the non-significant differences between the WMT and RT groups on Reading-chains is that the RT intervention improves the performance on the skills related to performance on the Reading-chains tests. The statistically significant improvement evident in both Letter- and Word-chains in the evaluation of outcome measures indicates that performance is enhanced even without any intervention. Since the participants in the earlier evaluation were undiagnosed, it is difficult to draw any conclusions on whether or not the gains in diagnosed children in the RT and WMT groups were due to test-retest effects or actual improved abilities.

This study had problems with high levels of drop-out. The rate of drop-outs was higher in the WMT group than in the RT group. WMT is an extensive intervention putting high demands on the children as well as the adults supervising the training. There are several possible explanations for children not completing training. For example conflicts between parent and child and lack of endurance could pose problems. In most Swedish families both parents work full time and it can be difficult to find the time needed for training in an already stressful everyday-life. Both the child and the parent supervising the training need to be fully motivated. The WMT program used in this study consisted of only 9 exercises that the child performed on each day of training. Some children may have found the lack of variation tiring.

It is possible that locating WMT to schools, with professional pedagogues functioning as supervisors instead of training with parents at home could diminish drop-outs. When WMT adds to the already high demands of school there is a risk that the total work load will be too much for the child to manage. Professional pedagogues are trained to motivate children and using pedagogues as supervisors would minimize role conflicts that can appear when parents act as supervisors to their own children. The effects on verbal working memory and mathematical reasoning motivate the extra efforts demanded by schools to offer WMT as an option in the school curriculum. Many of the children in this study already receive extra support and attention from pedagogues and school staff who could serve as supervisors for the training. Therefore the economic costs would not necessarily constitute a problem in introducing WMT in school. It is however possible that the rated decreases of ADHD-symptomatology reflect a positive effect of parent-child interaction that would be lost if WMT was located to schools. In one previous study children training both at home and in school were included (Klingberg, et al., 2005). No differences based on training location were found on the outcome measures in the study. Possible connections between drop-out rates and training location and the number of participants training at home and in school was not accounted for. Effects of training location must be further investigated to draw any certain conclusions on where training should be located.

Even though the present study did not aim at explaining the underlying mechanisms of ADHD, the results have implications on ADHD theories. The fact that effects on trained tasks in the WMT condition and gains in verbal working memory did not lead to significantly greater decreases of behavioral symptoms in the WMT condition than in
the RT condition, corresponds better to multiple pathways theories of ADHD (Sergeant, 2000; Sonuga-Barke, 2002) than with Barkley’s (1997) theory focusing on executive functioning as a primary deficit. If working memory deficits would account for all ADHD symptomatology, larger effects following WMT would be expected.

Since working memory deficits is not evident in all cases of ADHD and since no significant differences were found between the groups concerning reduction of ADHD symptoms, there may be better grounds for selecting WMT participants than the presence of an ADHD diagnosis. It is likely that assessing working memory and offering WMT to those with poor working memory is a better method.

Limits of the study
There are a few limits to this study which are of interest to mention. The first limitation regards experimental control. This study strived for a high ecological validity. The generalizability of the results is considerable since the sample in this study quite well resembles the actual seekers of care for ADHD related problems. The downside of aiming at ecological validity is partial loss of experimental control. All children trained in their home environment and such aspects as distractors, to which degree parents participated in the training and the amount and quality of feedback from the supervising parent during training could thus not be fully controlled. The effects of such aspects as mentioned should be evenly distributed across the interventions and not have any significant effects on the comparison of the groups.

The second limit of this study is that the training time per training occasion differs in the two interventions. The number of occasions for training was the same in both groups, but the WMT group trained longer periods of time per training day. The RT group showed no or marginal gains on all outcome measures except parent rating scales. It is highly unlikely that substantially larger gains in the RT condition would occur by simply increasing the amount of time used for training. If this was the case, tendencies towards improvement should have been apparent in the present study. The RT is not primarily designed for intensive training and the amount of words in the program would have made the training to repetitive if longer training time per day was required. The benefits of equal training time were not considered more valuable than the risk of extensive drop-outs in the RT.

The modest number of participants constitutes the third limitation of this study. Significant results might not be expected on some outcome measures with small effect sizes and small groups. The entire body of data from the tests, except delayed recall part of the AVLT, changed in a direction in favor of the WMT condition. It is possible that some actual gains did not yield statistical differences as a result of the low number of participants.

Future research
Other WMT studies have also had difficulties with relatively high drop-out rates (Klingberg, et al., 2005). Therefore it is essential to maximize the appeal of software aiming at training working memory for children. The external reward system seems to be an important aspect of the training but not sufficient to ensure compliance. The rewards included in the computer software could be further improved. More variation
regarding training exercises would also be preferable. The function of the adult supervising the training could be better standardized and evaluated.

So far studies on WMT have mainly compared the effects of WMT and control conditions without any expected effect. One study (Holmes, Gathercole, Place, Dunning, Hilton, & Elliott, 2009) also included effects of medication but did not make a direct comparison. Future research should focus on the differences between WMT and other commonly used interventions for the ADHD-population.

Future research should also investigate how much training is needed to obtain maximum improvements. One study (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008) compared different number of days of training, but had 19 days as maximum dose of training. It remains unclear if further training would have resulted in greater improvements. Another aspect of training dose that has not been thoroughly investigated is how long a training session should be to optimize the effect of WMT. As little as 15 minutes per day (Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2008) and as much as about 40 minutes of training (Klingberg, et al., 2005) has shown results.

So far most studies of WMT have not explored long term effects. Follow-up assessments have at the most been conducted 6 months post training for children (Holmes, Gathercole, & Dunning, 2009) and 18 months for adults (Dahlin, Nyberg, Bäckman, & Stigsdotter Neely, 2008). More studies are needed to further explore the long term effects on WMT.

Another aspect of WMT that needs to be further examined is the effects of different types of exercises. Preliminary results indicate that visuo-spatial exercises are superior to verbal exercises (Lucas, et al., 2008). Also, Thorell et al. (2008) found effects on verbal working memory following training with only visuo-spatial WMT exercises. The research on this subject has as of yet been insufficient to draw any certain conclusions. The effectiveness of different types of visuo-spatial and verbal exercises may also differ. The n-back task used by Jaeggi et al. (2008) is quite different to those used in this and other studies. No study to this date has addressed the question of exercise design.

Focus of research on WMT for children has mainly been on either the ADHD-population (Klingberg, et al., 2005) or children with poor working memory capacity (Holmes, Gathercole, Place, Dunning, Hilton, & Elliott, 2009). Since not all children diagnosed with ADHD have working memory deficits, it would be of interest to further investigate the effectiveness of WMT in different subgroups of the ADHD-population. Previous research indicates that working memory capacity can be improved in a normal population of children (Backman & Truedsson, 2008) and adults (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). No study has compared the effect size of WMT for participants with and without working memory deficits within a single study. It would also be of interest to examine whether WMT is as effective in the adult ADHD-population as in the child ADHD-population. There are also other psychiatric disorders involving working memory deficits, such as dyslexia (Smith-Spark & Fisk, 2007), and it would be of interest to explore possible gains of WMT in psychiatric disorders associated with working memory deficits.
It is noteworthy that many of the authors of earlier published WMT studies have had connections to the software developers. Studies performed by completely independent researchers would give stronger credibility to the WMT research field.

**Conclusion**
The result of the present study indicate that computerized WMT may lead to improvements of mathematical reasoning and verbal working memory in children diagnosed with ADHD. WMT did not result in significantly greater effects than RT on other measures of scholastic skills, visuo-spatial working memory or symptoms of inattention or hyperactivity/impulsivity. Studies on WMT with larger samples are needed to further investigate the effects of WMT.

**Acknowledgements**
The authors wish to thank Göran Söderlund, PhD, for supervising the project, Erik Truedsson, for insightful comments and funding of material and Richard Hultgren, for programming of tests. The authors also wish to thank Maria Södergren, for helpful comments on the manuscript. Finally Magnus Ivarsson thanks Eva-Maria Häusner and Stefan Strohmayer wishes to thank his wife, Linda and his children Ida, Elin, Lovisa, Anton and Emelie for their support and patience throughout the project.

**Disclosure**
Stefan Strohmayer was involved in the development of Memory Games. Magnus Ivarsson and Göran Söderlund have no financial relationships to disclose.
References


