

Working Memory Training and the Effect on Mathematical Achievement in Children with Attention Deficits and Special Needs

Karin I. E. Dahlin¹

¹ Department of Special Education, Stockholm University, Stockholm, Sweden

Correspondence: Karin Dahlin, Department of Special Education, Stockholm University, Stockholm, Sweden.

Tel: 46-708-679-122. E-mail: Karin.dahlin@specped.su.se

Received: October 20, 2012

Accepted: November 8, 2012

Online Published: January 16, 2013

doi:10.5539/jel.v2n1p118

URL: <http://dx.doi.org/10.5539/jel.v2n1p118>

Abstract

Working Memory (WM) has a central role in learning. It is suggested to be malleable and is considered necessary for several aspects of mathematical functioning. This study investigated whether work with an interactive computerised working memory training programme at school could affect the mathematical performance of young children. Fifty-seven children with attention deficits participated in an intervention programme. The treatment group trained daily, for 30-40 min. at school for five weeks, while the control group did not get any extra training. Looking at the group as a whole, mathematical performance improved in the treatment group compared with the control group directly following the five weeks of training (Time 2), but the results of the second post-test (Time 3, approximately seven months later) were no longer significant. Since there was only a small number of girls, the results were analysed for boys only. The boys had improved their mathematical results in both post-tests. WM-measures improved at Time 2 and 3 relative to Time 1 (pre-test) for the whole group, and for boys. Differences in training scores were related to differences in the non-verbal WM-measure Span board back.

The results indicate that boys aged 9 to 12 with special needs may benefit, over time, from WM training, as shown in the enhanced results in mathematics following WM training. However, as the intervention and control groups were not randomised, the results cannot be generalised; the results must be considered with caution.

Keywords: working memory training (WM training), attention, mathematics, special needs

1. Introduction

1.1 Working Memory

WM has a central role in learning and thinking and is conceptualised as the main cognitive system that stores and processes information. It is suggested that in order to remember information, it must first be processed in WM (Cowan, 2005). WM supports learning through the abilities to focus on the task in hand, inhibit irrelevant information and integrate information from several sources, including long-term memory (LTM). WM ability governs how successful the learning will be, as it is involved in processes necessary for achieving automatised knowledge (Cowan, 2005; Dehn, 2008).

Several theories describe the processing functions of WM. An “embedded system” is proposed by Cowan (2005), who states that attention control determines the outcome of cognitive processing. Ericsson & Kintsch (1995) propose a system that includes processing in LTM as well as in WM whilst performing skilled activities. This suggested long-term WM (LT WM) makes it possible to expand the processing and storing of specific tasks in WM (e.g., remembering many digits). LT WM is, for example, used in various professions, including by waiters (remembering and updating orders), doctors (recalling knowledge and identifying the correct medical diagnosis) and chess players (planning the next moves and considering the consequences of these).

The revised 1970s model presented by Baddeley & Hitch is often used in educational research. It consists of a central executive with three subsystems: the phonological loop, the visuo-spatial sketch pad and the episodic buffer (Baddeley, 2000). The central executive function (CEF), with its attention-control system, is considered the most complex component of WM. It eliminates unimportant information, coordinates ongoing processes of information, and controls strategies and inhibition, one of the most important functions of the CEF (Dehn, 2008). In the phonological loop, sound and speech are stored for a few seconds. The loop includes a “rehearsal”

function, the articulatory loop, which prevents the information from decaying, supporting loop storage capacity and verbal processing (Baddeley, 2000). The visuo-spatial sketch pad holds visual and spatial information. The episodic buffer interfaces with LTM and the CEF in WM, and may also control the awareness of consciousness (Baddeley, 2000). Baddeley's model, with the central executive and the three subsystems, is useful when trying to explain learning outcomes in relation to WM as it is often used in educational research.

WM is considered one of several executive functions (cognitive processes, EFs). EFs are described in various ways. Barkley (1997) synthesised a model of EFs from several theories, proposing that four EFs, which are necessary for behavioural inhibition, affect the motor control ability. One such function is WM. The other three functions are self-regulation (e.g., emotional self-control and regulation of arousal), internalisation of speech and reconstitution (e.g., analysis of behaviour). Children with attention deficit hyperactivity disorder (ADHD) have, for example, been found to be more active, more talkative, to make more errors and to have problems with inhibition compared with children without ADHD (Barkley, 1997).

1.2 WM and Attention Deficits/ADHD

WM makes it possible to concentrate and inhibit inappropriate information (Cowan, 2005). Attention difficulties seem to be more associated with deficiencies in cognitive development and with low school performance than with problems of hyperactivity and impulse control. Children with severe attention deficits, may be those who suffer significantly from visuo-spatial WM deficits (Castellanos & Tannock, 2002). Some of these children have ADHD. ADHD seems to be related to WM skills (e.g., Willcut, Doyle, Nigg, Faraone, & Pennington, 2005; Gropper & Tannock, 2009). However, children with ADHD are not a homogeneous group. Children with ADHD-I (inattention predominates) may sometimes be harder to identify than those with ADHD-C (ADHD-combined, attention and hyperactivity problems) and comorbidity may vitiate comparisons among children with ADHD, e.g., anxiety, learning disabilities, below-average language skills, conduct disorder (CD) and oppositional defiant disorder (ODD) (Nigg, 2006). Gender differences in ADHD have also been found: e.g., girls are less hyperactive and less impulsive. Moreover, more boys than girls are diagnosed with ADHD. The proportion of boys and girls with attention deficits ranges from 3:1 to 9:1, resulting in fewer girls than boys being examined in empirical studies (Gershon, 2002).

Gropper & Tannock (2009) found that university students with ADHD performed worse than the controls on one non-verbal measure, i.e., the Spatial span back. On the other hand, they were impaired on three out of four verbal WM measures. Willcut et al. (2005) found several executive functions to be associated with ADHD, such as planning, inhibition, attention and spatial WM abilities, which may negatively influence social interactions and academic outcomes (Miller & Hinshaw, 2010).

1.3 WM and Mathematics

Academic outcomes such as mathematics seem to be linked to WM (e.g., Swanson, 2006). Some children show difficulties in developing LTM representations of number concepts. It has been suggested that this may be caused by WM deficits (Geary, 1993). The importance of interactions between LTM and WM is also stressed (Gathercole & Alloway, 2008; Ericsson & Kintch, 1995). It has been found that number skills in kindergarten can predict skills in arithmetic in Grade 2 (Locuniak & Jordan, 2008) and that children with mathematical problems in Grade 5 still had problems one year later (Passolunghi & Siegel, 2004). Basic number skills problems in nine year old children are suggested to be related to counting problems at younger ages. Even simple a task such as the addition of pairs of single digits is a complex task (Hulme & Snowling, 2009).

WM is considered necessary for several aspects of mathematical functioning. Visuo-spatial WM, in particular, seems to be involved in children's ability to develop mathematical skills (Swanson, 2006). Gathercole, Alloway, Willis & Adams (2005) found that children needing special education with deficits in reading, also performed poorly in maths, visuo-spatial tasks and more complex memory tasks (such as repeating digits backwards).

Passolunghi & Siegel (2004) suggest that children with difficulties in solving word mathematical problems have a general deficit in the central executive function (Baddeley's model) and that they have a persistent deficit in WM that is not restricted to numerical WM tasks. Mathematical problems are related to deficit in the ability to inhibit unimportant information when attending to a task, and are also related to the ability to coordinate verbal and numerical information, and to understand concepts such as "smaller/larger" when comparing numbers (Passolunghi & Siegel, 2004). In a study using fMRI (functional Magnetic Resonance Imaging) analysis, Dehaene, Spelke, Pinel, Stanescu & Tsivkin (1999) found that exact calculations depend on language (number facts are stored in the same areas as language (e.g., days of the week), while approximation relies on non-verbal functions. This may explain why mathematical performance seemingly relies on phonological and visuo-spatial functions as well as on executive functions. Further, other variables also affect learning, such as strategies and

personality (e.g., Das, Naglieri, & Murphy, 1995).

1.4 Personality and Strategies

According to Das et al., intelligence and personality variables affect cognitive processes (the PASS theory: planning, attention and the simultaneous and successive coding of information) (1995). Intelligence can be described as being “the conscious capability of thought” (Escultura, 2012, p. 52). It includes memory, learning and the two most important skills for mathematics, i.e., critical thinking and creativity (e.g., reasoning, building new concepts and theories), formed by, for example, experiences, formal training, self-training and rational thought (Escultura, 2012).

The “planning ability”, as described by Das and colleagues (1995), may affect learning in certain situations in some children but not necessarily in others, despite there being no differences in children’s cognitive level. They found that children with adequate intelligence differ in problem solving skills. They suggest, in line with Escultura (2012), that these differences are to do with an individual’s personality and character, and are influenced not only by interactions between task demands and motivation, but also by predispositions and improvements in response to certain strategies and situations. It has been found that certain strategies affect memory performance (Carretti, Borella, & De Beni, 2007; Ericsson & Kintch, 1995).

1.5 WM Training

It is suggested that WM is malleable (Klingberg, Forssberg & Westerberg, 2002; Thorell, Lindqvist, Bergman, Bohlin, & Klingberg, 2008), and that some people can use WM capacity in a more effective way than others (Ericsson & Kintsch, 1995). In the Klingberg et al. study, functional MRI was used to measure brain activity during the performance of WM tasks. Four girls and nine boys (ages 9.4–18.5 years) participated. WM activity was noted in specific areas, and the older the children, the higher the activity was. WM capacity and activity in the same areas were also found to be related (Klingberg et al., 2002).

WM training studies target young children as well as adults with different problems (e.g., Thorell et al., 2008; Klingberg, Fernell, Olesen, Johnson, Gustafsson, Dahlström, et al., 2005; Jaeggi, Buschkuhl, Jonides & Perrig, 2007; Caviola, Mammarella, Cordonaldi & Lucangeli, 2009; Carretti, et al., 2007). Even if the conditions for these studies were not equivalent, they each draw the conclusion that WM training seems to improve results in young children as well as in adults with various profiles. In addition, strategy training positively affects memory performance (as suggested by Ericsson & Kintch, 1995) in adults as well as in children. Individual differences in WM may partly depend on a higher efficiency in WM processes when recalling items (Carretti et al., 2007). Treatment groups were taught how to use mental images when trying to remember 10–15 words on a list, and were also asked to orally express the quality of their images. In this case, the LTM is used for remembering the words and thereby supports more effective processes in WM. The control group was only asked to remember the words and recall as many as possible (Carretti et al., 2007). Both young and older people showed an enhanced ability in remembering words compared with the control group.

Previous studies in WM training in Sweden (Klingberg et al., 2005) were clinical and included children diagnosed with ADHD. Attention capacity and mathematical development seem to be related to WM abilities and WM seems to be impaired in children with ADHD / attention deficits. Therefore, WM training may be a promising intervention for students with attention deficits and mathematical problems. However, few studies have investigated both WM-training and mathematics. To address this deficit, this study considers the role of WM training in mathematics. It focuses on children with special educational needs who were educated in small groups in ordinary school settings, in separate rooms, either in the school building or nearby.

1.6 Aim

The aim of this study is to investigate whether a computerised WM-training program will influence WM and mathematical results in young children with attention deficits and special needs. The treatment group trained for five weeks with an interactive computerised training programme (cf. Klingberg et al., 2005) and underwent measures in mathematics and neuropsychological tasks before and after the intervention (directly following the training and seven months later). A control group completed basic skill measures within the same periods as the treatment group. Results at post-tests relative to pre-test were analysed.

It was hypothesised that WM training at school for a period of five weeks would improve skills in WM ability, and subsequently improve results in mathematics. The questions to be answered:

- a) How do children with attention deficits perform in mathematics after five weeks of working memory training at post-tests (directly following the training and seven months later) compared to the control group, who did not receive any extra training?

- b) How do children in the treatment group perform in WM-measures at post-tests compared to pre-tests?
- c) How are the outcome scores in WM measures, WM training results and mathematics related?
- d) Do boys and girls perform differently in WM-measures and/or in mathematics?

2. Method

2.1 Participant Characteristics

Children, in Grades 3 to 5, ($n = 57$, mean age = 10.7) participated in this study. They were all in regular school settings in Stockholm, Sweden, and neighbouring areas. The children, who were being educated in small groups, had attention deficits and special educational needs. Forty-two children constituted the treatment group (7 female) and 15 (4 female) made up the control group.

The following participation criteria were applied: (1) age 9–12 (Grades 3 to 5); (2) educated in small groups aimed at children with attention deficits; (3) earlier diagnosis of ADHD (by a doctor or psychologist), or attention deficits assessed by either school psychologists, to warrant small-group placement, or teachers and, in turn, by parent interviews with a psychologist (40 minutes); (4) Swedish as the individual's first language (teacher information); and finally, (5) absence of ODD (Oppositional Defiant Disorder; parent rating scales), mental retardation or autism (teacher/headmaster information). Seventeen children in the treatment group had received an ADHD diagnosis before joining the project, as had three members of the control group.

2.2 Sampling Procedures

An inventory of small classes (approximately 2–10 children in each class) of children with attention problems in Stockholm and neighbouring areas, acted as the frame for selecting classes and children for the project. These schools were invited to participate approximately five months before the intervention study was scheduled to begin. The schools were selected for the study according to the order in which we received the signed agreements we had requested, but no more than one class from each school district / community was accepted in order to ensure that different areas around Stockholm were represented. The 57 children were enrolled in 16 different schools, with two to five children in each class taking part in the study. Forty-two children from nine schools formed the treatment group in a first phase. Fifteen children from seven schools then formed the control group.

The number of participating schools was high because there is generally no more than one small group of qualifying children in each school, which includes a few children with a broad range of difficulties. For this reason, not all of the children in each class met the criteria for participation. The overall reason for specifically placing a child in a small class is that the child requires significant assistance, more than can be provided in a regular classroom.

2.3 Measures

2.3.1 Mathematics

Basic number skills Several number skills were measured with the Basic Number Screening Test (BNST) (Gillham & Hesse 2001). It is a nonverbal screening test with no time limit, suitable for children aged 7 to 12 years. The test uses 15 “number concept” items (e.g., place value, grouping) and 15 “number operations” items (i.e., carrying out tasks involving calculation). Examples of items include:

- *Calculations*, e.g., $36+24=\square$ $7\times 5=\square$ $56-42=\square$ (Instructions: “Here are three different calculations. (Pause) Look carefully at the signs and write the answers inside the boxes.”)

- *Patterns in series*, e.g., 37, 38, 39, $\square$ $\square$ (Instructions: “First you have three numbers and two empty boxes. (Pause) You put the **two** numbers that come **next** in the empty boxes.”)

- *Place-value*, e.g., 186 $\square$ (Instructions: “First you have a number beside an empty box. (Pause) In the box write the digit which stands for the tens.”)

- *Grouping* (Instructions: “There you have a large box with a lot of trees in it. (Pause) You draw lines round them to put them into **groups of seven** - then write the number of trees **left over** in the small box at the end.”)

- *Division* (Instructions: “Next you have the drawing of a bar of chocolate. (Pause) Suppose your mother says that you can break off a quarter of it to eat. Shade in a **quarter** of the chocolate bar to show what you've eaten”) (pages 9 & 10, the Manual, Gillham & Hesse, 2001).

The maximum possible score is 30. The test is designed for children from 7 years, and it therefore starts at an easy level, at page 1 and 2 (Grades 1 to 3), becoming progressively more demanding as the child works through

the pages to the last page (Grades 4 to 5). When three items in a row were solved incorrectly, the test session was closed.

Test scores are standardised and converted into number age norms for English children. To ensure that test norms could be transferred to Swedish children of the same ages, 100 Swedish children (51 male and 49 female) in regular classes in Grades 3, 4 and 5 sat the test once. Their mean number age was 10.25 years old. Their chronological mean age was 10.5. This result was similar to that of English children of the same age. There were no gender differences (male: $n = 51$, mean = 19.54 and female: $n = 49$, mean = 20.54) (t -test: $t(98) = -1.014$, mean difference = 0.9815, $p > .05$). Moreover, Pearson's t -test showed that Swedish children in regular classes performed significantly better in all mathematical measures compared with the children in the treatment and control groups ($p < .05$).

Addition and subtraction skills Addition and subtraction verification tasks were used to measure specific arithmetic skills. The children had two minutes to decide whether the given answers were correct by putting a cross (x) in a box after each item to indicate that it was correct, or a minus sign (-) if it was wrong (for example, $2 + 5 = 8$ ☐). The children were encouraged to try to solve as many items as possible. The tests were graded by adding one point per correct answer and subtracting one point for each error. In the first 15 items, only single digits were used with sums up to 15 (e.g., $5 + 3 = 7$); in the next five items, numbers with 1 or 2 digits were used (e.g., $8 + 14 = 22$) (pages 1 & 2). Thus, the test starts at an easy level and becomes progressively more demanding at pages 3 and 4, with sums up to 100 (e.g., $31 + 25 = 56$). There are four 3-digit items at the end (e.g., $123 + 255 = 378$). The subtraction test was designed the same way. The maximum attainable scores are 44 for addition and 40 for subtraction.

Parallel test versions The BNST manual gives the correlation between the two forms as $r = .93$. To ensure the reliability of parallel test versions, the Grade 3 students completed the addition and subtraction measures one week apart: addition ($r = .80$) and subtraction ($r = .83$).

2.3.2 WM Measures and Non-verbal Reasoning

Verbal WM 'Digit span' is part of WISC III (the Wechsler Intelligence Scale for Children). First, the child is asked to repeat verbally presented series of numbers in the same order as presented (this test is suggested to tap the phonological loop, Swanson, 2006). Second, the child is asked to repeat a verbally presented series of numbers in reverse order (this test is suggested to tap the CEF in WM, Swanson, 2006).

Visuo-spatial WM "Span board" is part of WAIS-NI (the Wechsler Adult Intelligence Scale). The administrator points at blocks on a bar. In the first part of the test, the child is asked to repeat the sequences in the same order, pointing to another set of blocks (this test is suggested to tap the visuo-spatial short-term memory). In the second part of the test, the child is asked to repeat the sequences in reverse order (loading onto visuo-spatial "sketch pad" and the CEF in WM (Swanson, 2006).

Non-verbal reasoning ability "Raven" (Raven's Coloured Progressive Matrices) is a non-verbal non-timed test. The child is asked to complete figural patterns by choosing the missing piece from six response options. The total score is 36.

2.3.3 Training Scores

The training scores obtained during working memory training sessions were converted into index scores and were used in the analysis, i.e., start index (the results from day one) and max index (the highest score achieved during the five weeks of training sessions).

2.4 The Training Programme

The children in the treatment group followed an interactive, computerised training program (Cogmed Working memory training) at school every day for five weeks. A fixed number of trials tapping verbal and visuo-spatial abilities was performed each day and completed in approximately 40 minutes. The training programme comprised eight different items, which were to be completed each day and totalled approximately 100 trials. Verbal WM tasks ($n = 2$) are to repeat a verbally presented series of numbers and a series of non-words in reverse order. Visuo-spatial WM tasks ($n = 3$) are, for example, to point out, in the opposite direction, the number of asteroids moving in the sky, lit up one by one. STM tasks ($n = 3$) to be completed are the repetition of visual and/or spoken information in the same order as it was presented, such as letters, syllabi or lights in specific positions. The level of difficulty was adapted to the WM capacity of each child through the training session (i.e., the increasing number of subjects to be repeated). Feedback was given immediately, verbally and visually, by scores indicated by a "thermometer" on a screen. An adult supported each child, one at a time, during the whole training period, which took place in a room next to the classroom. The parents received a daily report about how

their children were performing, which was signed by the adult responsible for the child's training sessions at school. It was taken home and brought back by the child every day. Once a week, a psychologist phoned the person responsible for each child's WM training, to give feedback and advice. The training scores were converted into index scores and were used in the analysis, i.e., start index and max index.

2.5 Procedure

The study is a quasi-experimental study with a pretest-treatment-post-test-post-test design. Three sessions of assessments were completed.

The treatment group received five weeks of computerised training of WM at school, daily for 30–40 minutes. The children underwent measures in mathematics and in neuropsychological tasks: (1) pre-test, Time 1; T1); (2) post-test, approximately six weeks later (Time 2; T2); and (3) post-test, approximately six to seven months later (Time 3; T3).

Fourteen children (2 female) from the treatment group (mean age = 10.5) also underwent 10 extra days of training before the second post-test (Time 3; T3). They were randomly selected and training was completed in the same way as during the first training session. A second post-test was then conducted (T3). The reason for this additional training was to investigate whether it would have any significant impact on the results at the second post-test for the 14 children completing this training compared to the rest.

The control group received regular special education training, and underwent the same basic skill measures within the same time intervals as the treatment group, but they did not complete measures in neuropsychological tasks due to insufficient economical resources.

2.6 Statistics and Data Analysis

Two types of models were used (SAS Institute Inc., 2004, version 9.1) to evaluate the treatment effect. Changes were modelled separately for T1–T2 and T1–T3 respectively. The main model type for estimating treatment effect (the repeated effects model) was a repeated effects mixed effects model, with individual measurement points as repeated random outcomes with separate bivariate normal distributions for the treatment and control groups respectively. Fixed effects were grouping (treatment/control) and time (T1/T2 and T1/T3 respectively), and fixed covariates were age and gender (the latter only in models for all children).

The other, complementary, model type (the gain score model) is a mixed effects gain score model. The gain scores are differences of outcomes at T2–T1 and T3–T1, respectively. When testing treatment effect, fixed effect was grouping (treatment/control), and when estimating baseline effect, fixed effect was grouping and outcome at T1. Fixed covariates were age and gender (the latter only in models for all children). The only random components here are individual errors in gain scores, with separate normal distributions for the treatment and control groups. Using a variation of the baseline effect model, the interaction between baseline effect and group (treatment/control) was included and tested.

The two model types give similar results when estimating treatment effects. The repeated measures model gave more flexible modelling possibilities, and was therefore preferred when estimating treatment effects. However, for studying baseline effects on increase, the gain score type is a suitable model, so therefore was used. (Note: by comparison, when estimating treatment effect, estimates at specific time points (T2 or T3), rather than change between time points, (T1–T2 or T1–T3) are controlled for baseline, T1).

All models estimate treatment effect as a specific effect of the training between T1 and T2. The effect of the extra training in parts of the treatment group between T2 and T3 is singled out as a separate estimate. Model results are shown for (all|boys) X (treatment effect | baseline effect on change) X (T1–T2 | T1–T3).

Cohen's *d*, a descriptive measure, was also calculated, comparing changes in the treatment group and the control group at post-tests with pre-tests:

$$\text{Cohen's } d = M1 - M2 / S \quad (1)$$

d is defined as the difference between the means, $M1 - M2$, divided by pooled standard deviation (mean *SD* of the two groups). This formula can also be used when comparing changes for one group (post-test minus pre-test). An effect size of $d = .20$ is considered small, $d = .50$ moderate and $d = .80$ large. This measure can have practical implications in school contexts and can be valuable in educational research, if used with caution (McMillan, 2004). Groups were also compared by *t*-tests. Finally, we conducted correlation analyses (Pearson correlation). We wanted to find out how changes in cognitive and training scores were related to mathematical measures.

3. Results

3.1 Baseline Data

As a basic check, the baseline results of the treatment and the control groups were compared. Standard test scores of the Basic Number Screening Test (BNST) are converted into number ages. There were no significant differences (*t*-test) either between the groups in number age at Time 1 ($p = 0.167$) or in pre-test mathematical scores (addition, $p = 0.232$, subtraction, $p = 0.248$, BNST, $p = 0.170$). Number ages for the treatment group were: T1 = 9.0 ($SD = 1.5$), T2 = 9.4 ($SD = 1.5$), T3 = 9.8 ($SD = 1.5$). Number ages for the control group were: T1 = 9.6 ($SD = 1.5$), T2 = 9.7 ($SD = 1.4$), T3 = 10.0 ($SD = 1.5$). The mean chronological age in the two groups was 10.7.

Next, we compared (*t*-test) two groups, equal in numbers ($n = 21$), within the treatment group ($n = 42$) using a median split: below and above median in number ages at T1: Group 1 = below median, Group 2 = above median. The mean number ages for Group 1 were: T1 = 7.8 ($SD = 0.8$), T3 = 8.7 ($SD = 1.2$). The mean number ages for Group 2 were: T1 = 10.2 ($SD = 0.9$), T3 = 10.45 ($SD = 0.9$). Group 1 performed significantly lower compared to Group 2 on all mathematics and WM measures except on Span board forward ($p = .72$, $t(39) = -1.1851$). Differences (T3 minus T1) in mathematics and WM measures for these two groups were compared (*t*-test). There were no significant differences between groups, i.e., both groups had changed their results as much at T3 compared to T1.

Finally, to find out how children in small groups perform in mathematics compared to children in regular classes, baseline results were compared. Pearson's *t*-test showed that Swedish children in regular classes performed significantly better in all mathematical measures compared with the children in the treatment and control groups ($p < .05$) (see 2.3.1). Available mathematical scores of regular classes (Grades 3, 4, 5) (BNST; $n = 134$, addition; $n = 99$, subtraction; $n = 116$) and results of the intervention groups (the treatment group and control group) ($n = 57$) at T1 were converted into z-scores. Mean z-scores for the intervention groups and regular classes could then be described and all three mathematical test results could be compared in between (Fig. 1).

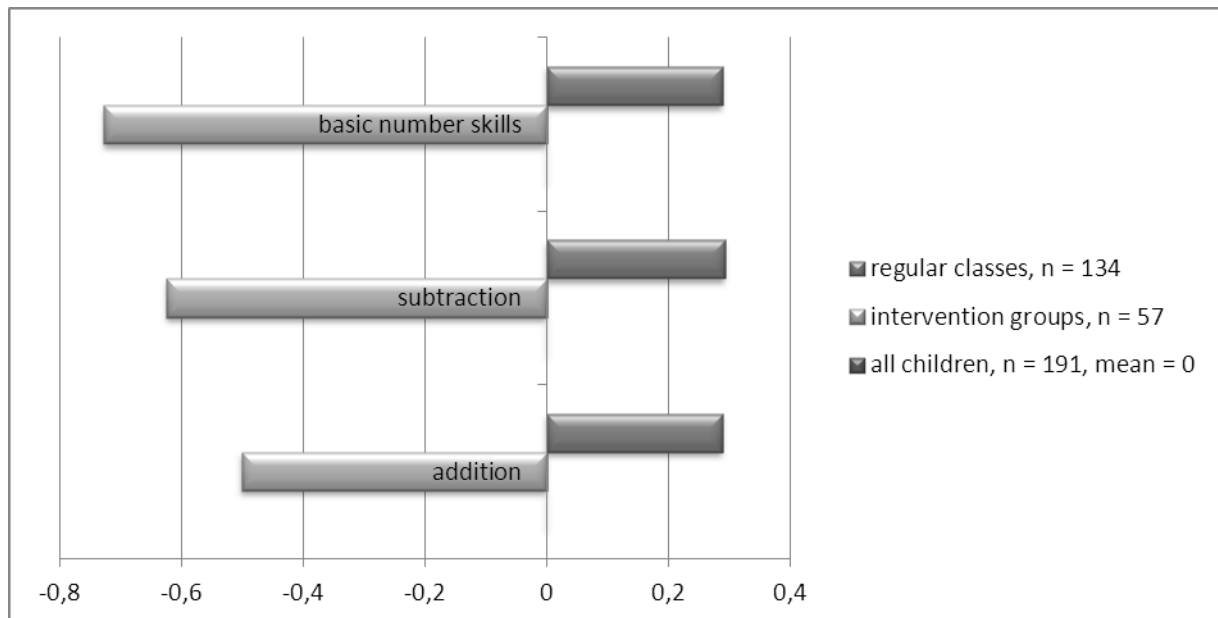


Figure 1. Mathematics z-scores in regular and small classes at Time 1

Figure 1 shows mean z-scores of the mathematics measures in the intervention groups, i.e., treatment and control groups ($n = 57$), and regular classes ($n = 134$) (Grades 3, 4, 5). Children in small classes perform below mean of the total group (all children) in all three measures. The lowest results occur in BNST (mean = -0.728, $SD = 1.05$) and subtraction (mean = -0.624, $SD = 0.94$). Addition mean scores = -0.501 ($SD = 0.96$). Children in regular classes perform equally in all three measures (mean = 0.29, $SD = 0.83$, 0.90). Pre-and post-test performances (mean scores) on the academic tests for the treatment group and the control group are reported in Table 1.

3.2 Working Memory Training Effects

The main question to be answered is whether mathematics results will improve following WM training. First, Cohen's d was calculated (Table 1).

Table 1. Descriptive statistics for mathematical measures in the treatment group and the control group (changes in the treatment group compared to changes in the control group [Cohen's d] are also reported)

	Treatment group (<i>n</i> = 42)						Control group (<i>n</i> = 15)						Treatment group compared with control group Cohen's <i>d</i>	
Time	T1		T2		T3		T1		T2		T3		T2–T1	T3–T1
	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>		
Addition	17.19	5.4	18.86	5.0	19.86	4.8	19.27	6.4	18.33	10.4	20.53	7.5	0.55	0.33
Subtraction	12.07	5.4	12.76	6.2	13.40	6.1	14.20	7.6	14.87	7.8	17.0	7.3	0.01	- 0.42
BNST	14.05	6.5	15.95	6.6	17.52	6.8	16.80	6.7	17.07	6.0	18.47	6.7	0.69	0.65

Note: SD = standard deviation; T1 = pre-test; T2 = post-test, approximately six weeks later; T3 = approximately seven months later; BNST= Basic Number Screening Test.

Table 1 shows pre-and post-test performances on the academic tests for all children in the treatment and control groups. The effect (d) on BNST (basic number skills) is large (0.69) compared to the control group directly after training (T2-T1), but six months later it is smaller (T3-T1).

Next, the treatment effect on BNST, addition and subtraction was estimated with the repeated measures model. BNST was significant at Time 2 (T2) (treatment effect = 1.638, $SE = 0.690$, $F(1, 26.5) = 5.63$, $p < .05$) but not at Time 3 (T3) (treatment effect = 1.690, $SE = 0.915$, $F(1, 27.6) = 3.41$, $p = .075$). The additional training of 10 days had no significant impact on the results at post-test for the 14 children completing this training (treatment effect = 0.36, $SE = 0.84$, $t(42) = 0.42$, $p = .67$).

The effect of baseline results at T1 on gain scores T2-T1 and T3-T1 was not significant (T2: effect = - 0.087, $SE = 0.050$, $F(56.9) = 2.96$, $p = .091$; T3: effect = - 0.035, $SE = 0.059$, $F(1, 55.2) = 0.36$, $p = .554$). Neither was the interaction between baseline effect and group significant (T2: $F(1, 28.9) = 1.30$, $p = .264$; T3: $F(1, 24.4) = 0.20$, $p = .661$). Therefore, there is no evidence that low and high performers would improve differently or that this pattern would differ between the control and treatment groups.

3.3 Gender Differences

We found that the girls ($n = 11$) performed differently compared with the boys (e.g., boys had higher number ages, BNST) compared with girls at T1, $p < .05$. Significant differences between boys and girls occurred within the treatment group in the WM Digit span forward measure, subtraction and the basic number test (BNST) at T1, T2 and T3, and in training scores, i.e., start index scores and max index scores). In contrast, improvements in WM training scores (mean scores: max index minus start index) were equal for boys and girls. However, the sample size for the girls was insufficient to test effects efficiently.

3.3.1 WM Training Effects: Boys

As girls performed differently compared with the boys, the analyses were repeated for the boys ($n = 46$) only. The treatment effect on BNST, addition and subtraction was estimated with the repeated measures model. BNST was significant at T2 (treatment effect = 1.891, $SE = 0.868$, $F(1, 17.7) = 4.74$, $p < .05$) as well as at T3 (treatment effect = 2.391, $SE = 0.999$, $F(1, 22.3) = 5.73$, $p < .05$). The additional training of 10 days had no significant

impact on the results at post-test for the 12 boys completing this training (treatment effect = 0.442, SE = 0.965, $F(1, 35) = 0.21, p = .650$).

The effect of baseline results at T1 on gain scores T2–T1 and T3–T1 was not significant (T2: effect = - 0.089, SE = 0.067, $F(1, 45.5) = 1.78, p = 0.189$; T3: effect = - 0.041, SE = 0.075, $F(1, 45.6) = 0.30, p = .589$). Neither was the interaction between baseline effect and group significant in boys (T2: $F(1, 20.1) = 0.73, p = .404$; T3: $F(1, 19.7) = 0.57, p = 0.454$). Therefore, there is no evidence that low and high performing boys would improve differently, or that this pattern would differ between the control and treatment groups. Effects were also calculated with Cohen's d (Table 2).

Table 2. Descriptive statistics for boys (changes in the treatment group compared to changes in the control group [Cohen's d] at T2 and T3 are also reported)

Time	Treatment group ($n = 35$)						Control group ($n = 11$)						Treatment group compared with control group Cohen's d	
	T1 <i>mean</i>	<i>SD</i>	T2 <i>mean</i>	<i>SD</i>	T3 <i>mean</i>	<i>SD</i>	T1 <i>mean</i>	<i>SD</i>	T2 <i>mean</i>	<i>SD</i>	T3 <i>mean</i>	<i>SD</i>	T2–T1	T3–T1
Addition	17.77	5.2	19.37	5.1	20.60	4.8	19.91	7.1	18.45	12.0	21.36	8.6	0.59	0.32
Subtraction	12.89	5.3	13.43	6.2	14.31	6.0	15.45	8.5	15.45	9.1	18.09	8.2	0.14	- 0.37
BNST	15.29	6.2	17.09	6.4	18.83	6.3	18.36	6.9	18.27	6.3	19.36	7.7	0.74	0.90

Note: *SD* = standard deviation; T1 = pre-test; T2 = post-test, approximately six weeks later; T3 = approximately seven months later; BNST= Basic Number Screening Test.

Table 2 shows pre and post-test performances for the academic tests in the treatment and control groups for boys. In the last two columns the Cohen's d effect is reported. As shown in Table 2 the effects (d) are large in BNST (basic number skills) for boys in the treatment group compared to boys in the control group (T2-T1, T3-T1).

3.4 Cognitive Measures within the Treatment Group

Due to insufficient economic resources, the control group completed neither WM-measures nor Raven. Therefore, it was not possible to compare results between groups. To find out if the treatment group and boys respectively improved in WM-measures and problem solving (Raven) at T2 and T3 relative to T1, effect sizes (Cohen's d) were conducted (Table 3, Table 4).

3.5 Correlations

Digit back and Spanboard back correlated with mathematics at T1, T2 and T3 ($p < .01$). Out of all of the WM measures, it was the Span board back that correlated most with mathematics in the treatment group at T1, T2 and T3 (e.g., Span board and BNST: T1, $r = .60$; T2, $r = .50, p < .01$, T3, $r = .49, p < .01$). All of the mathematics tests correlated (e.g., at T1: addition and subtraction, $r = .77, p < .01$; BNST and addition, $r = .70, p < .01$, BNST and subtraction $r = .73, p < .01$).

The outcome scores of training index (max index minus start index) for boys were significantly related to the outcome scores of Span board back (T3 - T1) ($r = .48, p < .01$). Span board back (T3-T1) was also related to addition results (T3 -T1) ($r = .41, p < .05$). The same pattern occurred for the total group (boys and girls) although the relations were slightly weaker. Path analysis would have been appropriate here to further investigate relationships, but the low number of participants in the present study made it less suitable, so the decision was made to forego these calculations.

3.6 Summary

In summary, for boys and girls as one group, the treatment effect on BNST was significant at T2 but not at T3. However, when analysing boys only, the treatment effect was significant at both T2 and T3. Addition and subtraction performances did not improve.

WM-measures were improved at T2 and T3 relative to T1. Span board forward and back improved at post-tests relative to pre-test for all children, while the short-term memory Digit forward did not improve greatly. Differences (T3 relative to T1) in training scores were related to differences in Span board back measures.

Table 3. Descriptive statistics (at Time 1, 2 and 3) for WM-measures and Raven in the total treatment group ($n = 42$) (differences between post-tests [T2, T3] and pre-test [T1] are reported [Cohen's d])

	T1		T2		T3		Cohen's d	Cohen's d
	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	T2-T1	T3-T1
Span board forward	4.54	0.9	5.68	1.0	5.55	1.0	1.19	1.05
Span board back	3.96	1.1	5.13	1.4	4.99	1.1	0.92	0.93
Digit span forward	4.24	0.7	4.72	0.9	4.63	0.9	0.59	0.48
Digit span back	3.02	0.7	3.71	1.0	3.34	1.1	0.79	0.34
Raven	27.23	5.0	29.68	4.7	29.90	4.5	0.50	0.56

Note: *SD* = standard deviation; T1 = pre-test; T2 = post-test, approximately six weeks later; T3 = approximately seven months later. Mean scores are reported in Dahlin (2011).

Table 3 shows that the effect (d) was large in the total group ($n = 42$) (T2-T1, T3-T1) on Span board forward and Span board back. The effects on Digit span were large, moderate or low. The effect on Raven was moderate. The same pattern occurred for boys (Table 4).

Table 4. Descriptive statistics (at Time 1, 2 and 3) for the results of WM-measures and Raven in boys ($n = 34$) (differences between post-tests [T2, T3] and pre-test [T1] are reported [Cohen's d])

	T1		T2		T3		Cohen's d	Cohen's d
	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	<i>mean</i>	<i>SD</i>	T2-T1	T3-T1
Span board forward	4.60	0.9	5.84	1.0	5.65	1.0	1.25	1.09
Span board back	4.10	1.1	5.32	1.2	5.09	1.0	1.05	0.93
Digit span forward	4.40	0.7	4.83	0.8	4.81	0.8	0.57	0.54
Digit span back	3.10	0.8	3.88	1.0	3.47	1.2	0.87	0.36
Raven	27.14	5.3	29.73	5.0	29.82	4.9	0.45	0.52

Note. *SD* = standard deviation; T1 = pre-test; T2 = post-test, approximately six weeks later; T3 = approximately seven months later.

As shown in Table 4, the effects (d) on boys ($n = 34$) of non-verbal WM-measures Span board forward and backwards were larger than the effect on Digit span. Digit back had not enhanced greatly at T3 according to these analyses. However, at T2 the effect was large.

4. Discussion and Conclusions

It was hypothesised that five-week WM training would improve skills in WM ability, and consequently improve

mathematical skills. We investigated whether mathematical skills improved in young children with attention deficits and special needs following interactive computerised WM training at school and how changes in WM-measures and training scores are related to mathematical outcomes in these children.

4.1 Mathematics

The treatment effect on the BNST (Basic Number Screening Test), addition and subtraction was estimated with the repeated measures model. No effects appeared for the addition or subtraction tests. The BNST was significant at Time 2 (T2) but not at Time 3 (T3) in the total treatment group (boys and girls). However, when analysing the boys' results only, the pattern differed from the analysis of mixed gender groups. The most surprising finding was that boys in the treatment group improved their BNST results at T2 (directly following the training) and at T3 (approximately seven months following the first post-test; T2), compared with boys in the control group, when the girls were excluded from the analysis.

This was an unexpected finding considering that the results from the mixed gender group showed a different pattern. For example, Holmes, Gathercole, & Dunning (2009) found no positive changes in mathematical reasoning directly after finishing five weeks of WM training (the same training programme as used in the present study), but found positive changes later, after six months. However, only the treatment group was re-tested. Holmes et al. claim that changes in skills take time to develop following a training period. In their study, there were almost as many girls as boys in the treatment group ($n = 22$, 10 female/10 male) while in the control group the girls were fewer ($n = 20$, 5 female/15 boys). In the present study, 1/6 were girls ($n = 42$, 7 female/35 boys) in the treatment group and 1/3 in the control group. This fact may have an impact on the results. At any rate, the proportion boys and girls clearly affected the results in the present study. Various kinds of measures are employed in studies, which obviously makes comparing studies problematic.

The mathematical test which showed improvement (BNST) contains non-verbal items and differs in many ways from the timed two-minute addition and subtraction verification tasks. The BNST comprises various items, both calculations when it is clear which kind of calculation to carry out, and items like grouping a number of trees in a box. Therefore, this test does not rely on automatic recall as much as the addition and subtraction tests do (cf. Dehaene et al., 1999), in which it is favorable to quickly collect the answer from LTM, during just two minutes. Also, in the BNST, clear instructions are read to the child by an adult, one item at a time, and the test is not time limited. These two factors may affect the ability to focus better on the task (cf. Das et al., 1995). As the instructions are read (twice if required) to the child in the BNST task, one item and instruction at a time, this probably helps the child to focus on each specific item. Obviously, it is worth discussing verbal instruction here, as the child has to understand the instructions and the specific terms. Children with no vocabulary problems could benefit from a test such as this. The design of the test may hence affect the results. At start (Time 1) this was the test that posed the most problems for the treatment and control groups compared to addition, subtraction and regular classes (see Fig. 1).

Further, knowledge about mathematical facts and rules, in cooperation with cognitive processes, such as attention and planning ability, may affect the results of various items, including not only adding and subtracting, but also items such as grouping, patterns and values, assessing the understanding of the number system (cf. reading comprehension problems which can depend on poor decoding (dyslexia) or, for example, poor vocabulary and grammar despite reading words accurately) (Hulme & Snowling, 2009). Mathematical problems are also related to the ability to understand mathematical vocabulary, as suggested by Passolunghi & Siegel (2004). Knowing that there is no time limit (as in the BNST test) may also be a positive factor for the child, rendering the situation less demanding and most likely increasing motivation (Das et al., 1995).

No effects appeared for the addition and subtraction tests, which are related to automatic recall and speed and probably rely more on LTM than do BNST. The tests focus on addition and subtraction, respectively, for two minutes. The importance of automatic recall and speed, and the suggestion that some people show difficulties in consolidating number concepts, are stressed (Geary, 1993). If simple addition and subtraction tasks are not consolidated the answer cannot be collected from the LTM quickly (Geary, 1993), or if there are no beneficial strategies (e.g., having to rely on counting up to the first digit, and then counting on). This takes time, even if the items are very simple (Hulme & Snowling, 2009) and there will be difficulties when trying to solve as many tasks as possible in a limited time. Further, WM may become overloaded and one may lose track (Clark, Nguyen, & Sweller, 2006). As a result, one scores low on tests of this kind.

4.2 Girls

In this study, girls performed at a lower baseline level compared to boys. However, since there was only a small number of girls in the study, no conclusions can be drawn from their results.

The question of whether girls with attention deficits generally perform differently from boys with attention deficits can only be answered by investigating a larger number of girls with attention deficits and special needs completing WM training, and by comparing their results to those of girls and boys with and without attention deficits.

Studies show that gender difference may occur in children with ADHD/attention deficits, e.g., girls with ADHD show persistent deficits in executive functioning and have a higher risk, compared with girls without ADHD, of developing educational and antisocial disorders (Biedermann, et al., 2008). Attention deficits in girls may go unrecognised, as they seem to be less hyperactive and less inattentive compared to boys, and girls may not get the specific help they need in order to cope positively (Gathercole & Alloway, 2008). One further argument about the girls' performance is that females may have to be more severely affected by ADHD to be identified, particularly by classroom teachers. However, the poor baseline results in these particular girls may simply be a matter of chance. The ratio of girls and boys was in fact unequal ($n = 57$: 11 female, 46 boys).

4.3 WM-Measures

Neither verbal short term memory (Digit forward, $d = 0.48$) nor verbal WM (Digit back, $d = 0.34$) (T3-T1) improved significantly in the treatment group. It has been argued that the phonological loop is important for storing the original addends when doing mathematics, and that counting speed relies mainly on the phonological loop, i.e., the articulatory loop (Passolunghi & Siegel, 2004), one of the subcomponents of Baddeley's WM model. The most improved WM-measures in the present study were Span board forward and back at Time 3 relative to Time 1 in the experimental group (boys and girls), and in boys.

Similar findings are reported from the Holmes, et al. study (2009) using the same WM training programme as in the present study: WM training had no effects on verbal STM, but it did have effects on visuo-spatial STM and verbal and visuo-spatial WM. It may be that STM (the phonological loop) is not affected by the training, while the CEF (central executive function) is.

This reasoning is in line with findings from other studies. Passolunghi & Siegel (2004) found that children with mathematical disabilities had a persistent weakness in WM, in the central executive functioning (CEF), and particularly in the WM task Digit span back measure. Swanson (2006) found that mathematical skills in young children with average and above-average mathematical scores rely more on CEFs than on the phonological loop. Swanson argues that the CEFs seem to play a critical role in mathematics, independently of the phonological loop and the visuo-spatial "sketch pad". He found that the CEF contributed approximately 12 % of the variance in maths calculations.

4.4 Explanations

The present study shows that computerised training improved untrained skills in mathematics, but it is still unclear how the training affected specific WM functions. We know that visuo-spatial WM measures improved in children in the present study. Neuroimaging studies report that increased brain activity was observed in the prefrontal cortex (areas associated with WM), following working memory training (Olesen, Westerberg, & Klingberg, 2004; Westerberg & Klingberg, 2007). In these studies the same training programme was used as in the present study. It is suggested that WM depends on these frontal and parietal areas (e.g., Olesen et al., 2004; Goswami, 2008).

As a result of increased brain activity, efficiency in processing and the ability to focus on the task in hand were perhaps enhanced. It is suggested that individual differences in WM may partly depend on a higher efficiency in WM processes (Carretti et al., 2007), and that inhibition control affects the outcome when solving mathematical problems (Passolunghi & Siegel, 2004). This may be one explanation.

Motivation, personality, certain strategies and situations also affect learning and are most important in all cognitive processes (Das, et al., 1995; Escultura, 2012). Responses to certain strategies and situations may vary in individuals. Interactions between task demands are also significant. The children in the treatment group may consider that they got paid for hard work during the training session (positive feed-back), and realise that they are able to manage if they try hard to focus on the task in hand, an awareness that they can apply at school.

However, learning is not easily explained. Various cognitive processes (EFs) seem to be important for mathematical development and for all learning (Barkley, 1997; Passolunghi & Siegel, 2004). One key piece is WM-ability, in particular the cooperation with subcomponents of WM, and LTM (Gathercole & Alloway, 2008; Ericsson & Kintch, 1995). Visuo-spatial WM ability improved in children in this study and these skills in particular seem to be involved in the children's ability to develop mathematical skills (Swanson, 2006).

4.5 Limitations

The limitations of the present study are the differences in age, group sizes and the selection of the children. Selection was based on attention problems as rated by teachers and psychologists, and on the schoolroom context, regardless of gender. However, the criteria for participation ensured that the selected children were alike in many ways: they required special education, were in Grades 3 to 5, were working in small groups and had attention deficits but no ODD, autism problems or intellectual disability. There were also differences in pre-test scores between the treatment and the control groups. Analysing differences in baseline results was controlled, by using age and gender covariates in the analysis. Finally, the training did not improve performance on all mathematical tasks or WM-tasks so it is therefore not very likely that improvements were due to increased reinforcement from adults. However, as the intervention and control groups were not randomized, the results cannot be generalised; the results must be considered with caution.

4.6 Conclusions

The results show that boys in the treatment group improved their results in the BNST (Basic Number Screening Test) at post-tests compared with boys in the control group. Cognitive measures improved at post-tests relative to pre-test for boys, and for the total group (boys and girls) in the treatment group. Training scores (max index- start index) were related to differences in Span board back measure, which in turn were related to addition, and all the mathematics tests were correlated with each other. Span board back results were highly correlated to mathematics at T1, T2 and T3. A conclusion is therefore that the training effect may be related to the specific training programme and not to a general improvement.

It seems that the training programme may yield rewards for boys with attention deficits and special educational needs, that WM capacity may be able to improve and that WM training has a positive effect, as shown in one of the mathematical measures. This is in line with the hypothesis, suggesting that memory capacity is malleable (e.g., Ericsson & Kintsch, 1995; Ericsson, 2010) and that WM training seems to affect untrained skills in children and adults (Caviola et al., 2009; Thorell et al., 2008; Klingberg, et al., 2005; Jaeggi et al., 2008; Holmes et al., 2009).

4.7 Practical Implications

This study has some practical implications. First, it seems clear that measures of WM capacity could be useful for the early identification of children at risk of failing at school: in the present study, children with attention deficits and special needs scored lower in mathematics than children from regular classes. Second, clear instructions, without unnecessary words, focusing on one item at a time (BNST), seem to be beneficial for children with attention deficits. Third, in this study, girls performed significantly lower compared to boys in subtraction and BNST. It is important to understand children's attention difficulties, attitudes to learning and the various needs that individuals may have depending on the nature and severity of their problems (various ADHD profiles, WM-ability, planning ability and personality). Goldstein & Naglieri (2008), for example, propose that children with ADHD require a different kind of intervention, depending on the presence (or lack) of cognitive weakness. Therefore, children with ADHD but without cognitive weaknesses are helped via support in improving behaviour and a changed environment, while those with ADHD and cognitive weaknesses require advanced academic instruction targeting the individual's specific problems.

It seems to be very important as a teacher to be familiar in WM functions because weaknesses in WM ability seem to affect school performance, underscoring the need for teaching that minimises WM overload (Clark et al., 2006). Obviously, diagnostic proceedings are insufficient for effective pedagogical interventions for all pupils. Individual variations may go unselected in the diagnostic procedure. Problems vary in individuals with attention deficits, not only according to the most common criteria in, for example, ADHD diagnosis, but also to cognitive status. WM capacity is one of many components that have an impact on learning. It is important to be aware of the complexity of learning problems: one intervention alone is not enough to gauge an individual's learning problems. For example, automatic recall when carrying out basic calculations is one of several skills necessary for mathematical development, (cf. decoding and reading comprehension) (Loucinak & Jordan, 2008). In fact, the BNST, i.e., basic number skills, improved after the intervention although addition and subtraction skills did not improve. In the BNST, it is necessary to complete different calculations and use various number concepts such as grouping and placing value (in contrast to addition and subtraction tests) (see 2.3.1). Finally, further studies are needed to investigate the effects of WM on mathematical achievements, in girls with attention deficits/ADHD in particular. There are many questions that remain answered in this area.

Acknowledgement

This research was funded by The Swedish Research Council. I wish to thank all of the children, parents and teachers who made this study possible. I wish to thank Anders Skarlind, Stockholm University, for help with statistical analysis. I would also like to thank Professor Torkel Klingberg, the Karolinska Institute, for all of his help throughout the study, and Professor Mats Myrberg and Professor Mara Westling Allodi, Stockholm University, for their valuable support and assistance. I take full responsibility for the accuracy of data. I have full access to data and analysis in the study.

References

- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, 4(11), 417-423.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, 121(1), 65-94.
- Biedermann, J., Petty, C. R., Doyle, A. E., Spencer, T., Henderson, C. S., Marion, B., Fried, R., & Faraone, S. V. (2008). Stability of executive function deficits in girls with ADHD: A prospective longitudinal followup study into adolescence. *Developmental Neuropsychology*, 33(1), 44-61. <http://dx.doi.org/10.1080/87565640701729755>
- Carretti, B., Borella, E., & De Beni, R. (2007). Does strategic memory training improve the working memory performance of younger and older adults? *Experimental Psychology*, 54(4), 311-320. <http://dx.doi.org/10.1027/1618-3169.54.4.311>
- Castellanos, X. F., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Neuroscience*, 3, 617-627. <http://dx.doi.org/10.1038/nrn896>
- Caviola, S., Mammarella, I. C., Cornoldi, C., & Lucangeli, D. (2009). A metacognitive visuospatial working memory training for children. *International Electronic Journal of Elementary Education*, 2(1), 122-136. Retrieved from http://wmlabs.psy.unipd.it/Publication/caviola/Caviola%20et%20al._2009_A%20metacognitive%20visuospatial%20working%20memory%20training%20for%20children.pdf
- Clark, R., Nguyen, F., & Sweller, J. (2005). *Efficiency in learning: evidence-based guidelines to manage cognitive load*. San Fransisco: John Wiley & Sons, Inc.
- Cogmed Working Memory training. (2005). Cogmed Medical System AB, Stockholm, Sweden. Retrieved from <http://www.pearsonassessment.se>
- Cowan, N. (2005). *The working memory capacity*. Sussex: Psychology Press.
- Dahlin, K. I. E. (2011). Effects of working memory training on reading in children with special needs. *Reading & Writing*, 24(4), 479-491. <http://dx.doi.org/10.1007/s11145-010-9238-y>
- Das, J. R., Naglieri, J. A., & Murphy, D. B. (1995). Individual differences in cognitive processes of planning: A personality variable? *Psychological Record*, 45(3), 1-13.
- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, 284, 970-974.
- Dehn, M. J. (2008). *Working memory and academic learning: Assessment and intervention*. USA, New Jersey: John Wiley & Sons.
- Ericsson, A. K., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245. <http://dx.doi.org/10.1037//0033-295X.102.211>
- Ericsson, K. A. (2010). Superior memory of experts and long-term memory. An updated and extracted version of Ericsson. Retrieved from <http://www.psy.fsu.edu/faculty/ericsson/ericsson.mem.exp.html>
- Escultura, E. E. (2012). The physics of intelligence. *Journal of Education and Learning*, 1(2), 51-64. <http://dx.doi.org/10.5539/jel.v1n2p51>
- Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A. M. (2005). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, 93(3), 265-281. <http://dx.doi.org/10.1016/j.jecp.2005.08.003>
- Gathercole, S., & Alloway, T. P. (2008). *Working memory and learning*. England, London: SAGE Publication Ltd.

- Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological and genetic components. *Psychological Bulletin*, 114(2), 345-362. <http://dx.doi.org/10.1037/0033-2909.114.2.345>
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorder*, 5(3), 143-154. <http://dx.doi.org/10.1177/108705470200500302>
- Gillham, B., & Hesse, K. (2001). *Basic Number Screening Test*. England, London: Hodder and Stoughton Educational.
- Goldstein, S., & Naglieri, J. A. (2008). The school neuropsychology of ADHD: Theory and assessment, and intervention. *Psychology in the Schools*, 45(9), 859-874. <http://dx.doi.org/10.1002/pits.20331>
- Goswami, U. (2008). *Cognitive development. The learning brain*. Hove and New York: Psychology Press.
- Gropper, R. J., & Tannock, R. (2009). A pilot study of working memory and academic achievement in college student with ADHD. *Journal of Attention Disorders*, 12(6), 574-581. <http://dx.doi.org/10.1177/1087054708320390>
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, 12(4), F9-F15. <http://dx.doi.org/10.1111/j.1467-7687.2009.00848.x>
- Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. UK: Wiley-Blackwell.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *PNAS (The National Academy of Sciences of the United States of America)*, 105(6), 829-833. http://dx.doi.org/10.1073_pnas.0801268105
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience*, 14(1), 1-10. <http://dx.doi.org/10.1162/089892902317205276>
- Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD – A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44(2), 177-186. <http://dx.doi.org/10.1097/00004583-200502000-00010>
- Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities*, 4(5), 451-459. <http://dx.doi.org/10.1177/0022219408321126>
- McMillan, J. H. (2004). *Educational Research. Fundamentals for the consumer*. USA: Pearson Education, Inc.
- Miller, M., & Hinshaw, S. P. (2010). Does childhood executive function predict adolescent functional outcomes in girls with ADHD? *Journal of Abnormal Child Psychology*, 38, 315-326. <http://dx.doi.org/10.1007/s10802-009-9369-2>
- Nigg, J. T. (2006). *What causes ADHD?* USA, New York: Guilford Press.
- Olesen, P., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, 7(1), 75-79. <http://dx.doi.org/10.1038/nn1165>
- Passolunghi, C. M., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology*, 88, 348-369. <http://dx.doi.org/10.1016/j.jecp.2004.04.002>
- Raven, J. C. (1995). *Coloured Progressive Matrices*. Oxford, United Kingdom: Oxford Psychologists Press.
- SAS Institute Inc. (2004). *SAS/STAT® 9.1 User's Guide*. Cary, NC: SAS Institute Inc.
- Swanson, H. L. (2006). Cognitive processes that underlie mathematical precociousness in young children. *Journal of Experimental Child Psychology*, 93, 239-264. <http://dx.doi.org/10.1016/j.jecp.2005.09.006>
- Thorell, L. B., Lindqvist, S., Bergman, S., Bohlin, G., & Klingberg, T. (2008). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 11(6), 969-976. <http://dx.doi.org/10.1111/j.1467-7687.2008.00745.x>
- Westerberg, H., & Klingberg, T. (2007). Changes in cortical activity after training of working memory – A single-subject analysis. *Physiology & Behavior*, 92, 186-192.

<http://dx.doi.org/10.1016/j.physbeh.2007.05.041>

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention deficit/hyperactivity disorder: A meta-analytic review. *Society of Biological Psychiatry*, 57, 1336-1346. <http://dx.doi.org/10.1016/j.biopsych.2005.02.006>