Is working memory capacity and metacognitive training effective in enhancing school based reasoning achievements?

A synthesis of the research

Roel J.F.J. Ariës, Joris Ghysels, Wim Groot, Henriëtte Maassen van den Brink
Is working memory capacity and metacognitive training effective in enhancing school based reasoning achievements? A synthesis of the research

Roel J.F.J. Ariës1  Joris Ghysels2  Wim Groot2  Henriette Maassen van den Brink2

Abstract:
Secondary school-pupils frequently underachieve in tests which require problem-solving skills. Training of working memory (WM) may improve problem solving skills. We review primary and secondary school-based interventions to evaluate the capability WM-training to enhance problem solving skills. Two aspects of WM appear to affect reasoning skills. WM capacity plays a role in short term storing and in manipulating information. Meta-cognitive WM concerns the storage of acquired knowledge of problem solving skills in long term memory. We conclude that both aspects could improve reasoning skills. Metacognition causes students to achieve significantly better compared with non-trained groups regardless of training type and intensity. Few studies are found that focus on school-based interventions. Furthermore, in line with earlier reviews, many studies lack a critical analysis of the mechanisms by which WM causes reasoning achievements in classroom environments. The question whether transfer-effects to domain-specific reasoning is expected remains a topic for further research.

Key words: problem solving skills, secondary education, reasoning, working memory

1 Corresponding author: roel.aries@maastrichtuniversity.nl
2 Top Institute for Evidence Based Education Research (TIER), TIER-Maastricht University, Kapoenstraat 2, 6211 KW Maastricht, the Netherlands
Introduction

Working memory (WM) training, the training of the brain function that reorganizes and manipulates retained information, may enhance a broad variety of academic skills (Au, Sheehan, Tsai, Duncan, Buschkuehl, Jaeggi, 2015). In recent years, research on working memory has provided evidence that WM training could improve problem solving skills by enhancing the ability to provide solid analyses and answers on novel problems, e.g. reasoning questions (Conway, Kane & Engle, 2003; Kramarski & Mevarach, 2003; Cheshire, Ball & Lewis, 2005; Jaeggi, Buschkuehl, Jonides & Perrig, 2008) and improve test results (Agin, 2001; St. Clair-Thompson, Stevens, Hunt, & Bolder, 2010). The evidence on this is mixed, however, as other (review) studies have reported no significant improvements caused by WM-training (e.g. Redick, Shipstead, Wiemers, Melby-Lervåg & Hulme, 2015). Also, the excitement about the prospects for far transfer, i.e. ‘performance benefits in outcome measures that are contextually, structurally and superficially dissimilar to the trained task’ (Beatty & Vartanian, 2015, p. 1), has resulted in studies that investigate the enhancement of reasoning with novel information by trained participants. This is based on the view that WM is the main brain function that is activated when reasoning takes place, and its (prospected) ability to reason and solve new problems, independent of any past knowledge. However, few WM training strategies have been specifically developed and evaluated to improve problem solving skills and reasoning test results in primary and secondary education (Kramarski & Mevarach, 2003; Mevarach & Kramarski, 2003; Jolles, De Groot, Van Benthem, Dekkers, De Glopper et al., 2006; Alloway & Gathercole, 2009; De Jong, Van Gog, Jenks, Manlove, Van Hell et al., 2009; Lee, Ng & Ng, 2009; Redick, Shipstead et al., 2015).

WM plays a crucial role in developing problem solving skills because it underlies several cognitive abilities, including logical reasoning and problem solving. In this regard, WM is the activation of cortical cognitive networks (cognits), with subcortical support, to selectively attend to recent information for a prospective decision, choice, or action to solve a problem or to reach a goal. Two aspects of WM affect the process of problem solving. (1) WM temporarily stores and manipulates information during complex cognitive activities, such as reasoning. Holmes, Gathercole and Dunning (2009) define WM as the cognitive system that provides temporary storage and manipulation of information in the course of complex cognitive activities. According to Olesen, Westerberg and Klingberg (2004), the amount of information that a person can retain is an important factor for problem solving: WM needs to change its content constantly and rapidly because problem solving abilities often require processing different types of information, swiftly switching from one to another cognitive task and simultaneously solving diverse problems (Goldberg, 2010). The capacity of WM could therefore empower reasoning abilities (Sweller 1988; Bull & Scerif 2001; Klingberg, Forssberg & Westerberg, 2002; Holmes et al., 2009). Adaptive training of WM or related executive functions lead to substantial and sustained enhancement of (initially poor) WM in children. (2) Prescriptive knowledge, the generic memories of effective ways to solve higher leveled cognitive problems, is stored in WM (Gazzaniga, Ivry, Mangun & Steven, 2009; Klingberg, 2009; Goldberg, 2010). Experts, unlike novices, possess generic memories-developed schemas or patterns to solve problems (Sweller, 1988). This ‘pattern recognition’ is based on learning processes which underlie repeated exposure to similar contexts that originate in matched response strategies. According to Goldberg (2010), a transition occurs
from the absence of effective behavior of working memory to the formation of effective behavior in reasoning processes to address reasoning questions more effectively and efficiently when meta-cognitive components are added to training strategies. Therefore, meta-cognitive training causes working memory to form blueprints and to plan for rational analysis and analytical methods’ (Gold, et. al., 1996; Gazzaniga, et. al., 2009; Goldberg, 2010).

Despite the lack of consensus about the effect of WM training, it is concerning that many of the existing studies and reviews (e.g. Shipstead, Redick & Engle 2012; Melby-Lervåg & Hulme, 2013; Titz & Karbach, 2014) on WM and reasoning cannot be replicated in compulsory education because of its specifics and deficits. Van der Sluis, De Jong and Van der Leij (2007) for example argue that in most studies involving children ‘the executive performance in samples of normal children is compared to that of children in clinical samples, e.g. children with learning difficulties’ (p.430) or ADHD and are correlational studies (e.g. Bull, Johnston & Roy, 1999; Bull & Scerif, 2001; Redick et. al. 2015). Furthermore, many experimental studies primarily use Fluid or General Intelligence tests; e.g. Shipstead et al. (2012) argue that the training and testing of basic cognitive functions is a far cry from training of actual school based reasoning skills. In a more recent review of the effects of WM training on a broad range of academic skills, Redick et. al. (2015) supports this claim by concluding that WM training alone produces ‘no advantage for academic or achievement-based reading and arithmetic outcomes’ (p.1). However, nearly all studies included in their review have used samples of students with learning difficulties, which could affect learning outcomes compared to normal learning behavior. Furthermore, studies increasingly indicate that the implementation of subject-oriented content and processes is key to enhancing WM training models (Vos 2001, Van der Sluis et al., 2007; Lee et al., 2009). Other, more recent studies suggest that training on WM tasks correlate with (short term) improvements in academic achievements in reasoning (e.g. Au, Sheehan et. al. 2015; Stephenson & Halpern, 2013). Moreover, many experimental and correlational studies have tested children in their early years and in late adolescence (18-21 years old), and are thus not focused on primary and/or secondary education. Early and middle adolescence is a period of high vulnerability to reward orientation, risk-taking and problems in regulations of affect and behavior, and cannot be compared to late adolescence (e.g. regulatory competence) according to Steinberg (2006). Our synthesis of the literature distinguishes itself from these previous reviews on school-based WM training by specifically focusing on achievement outcomes with students with normal learning behavior as well as specifically focusing on two aspects of WM that are suggested to improve reasoning achievements.

This review addresses these issues by reviewing studies that focus on primary and secondary education to analyze whether WM training is effective in education settings to develop problem solving skills. The search strategy on which this review is based is focused on school-based interventions, as well as school tests.

This paper proceeds as follows. First the search strategy is discussed. The following section presents the findings. The last section discusses and concludes.

Method and search strategy

This paper focusses on school-based working memory capacity interventions as well as (metacognitive) reasoning strategy interventions. We only include articles that directly link
working memory training interventions to the children’s educational problem solving achievements. The review uses peer-reviewed journal articles from 1990 to 2015. Journal articles have been selected via databases ERIC (Educational Resources Information Center) and Psycinfo (see Figure 1 for flowchart). Google Scholar has been consulted to search for more articles.

The following keywords have been used as key search values: working memory (capacity), (primary/secondary) education, reasoning strategies, metacognition, school-based. The combination of at least three of these keywords yielded 145 articles. We found that many articles still did not address the school-based and content-based approach that we look for. Therefore, we added the following search terms to address the school-based approach more specifically: adolescence, young adults, executive function training, reasoning skills, critical thinking, cognitive learning, brain based learning, academic achievement, compulsory education, content based. The search strategy yielded 73 articles. We extracted 51 articles based on the following inclusion criteria: English or Dutch language, participants in randomized groups between five and twenty years old, publication year 1990-2015.

The greater part of the literature describes school based interventions, but did not use school tests to measure improvements in reasoning skills. In fact, they generally use general intelligence tests or tests for general reasoning. Also, many are correlational studies which do not allow for causal inferences. Although WM capacity and general intelligence are generally acknowledged as strong predictors of academic achievements, we exclude studies that use these general intelligence tests (16 articles). This is due to the absence of strong correlations between fluid intelligence training tasks (which is the capacity to reason and solve new problems, independent of any past knowledge) and school based reasoning, as far transfer of WM training to content based reasoning abilities seems very limited (e.g. Titz & Karbach, 2014). Correlational studies and tests between normal children and children with learning difficulties are also not included in this review (19 articles) as we focus on training and normal learning behavior. We have found a total of 15 articles (10 articles concerning working memory capacity training, 6 articles concerning metacognitive training) eligible for inclusion in the systematic review (tables 1 and 2). The selected articles have been published in peer reviewed journals and used (semi-)randomized groups of children aged five to seventeen.

Results

Current consensus in neuroscientific research is that developing a problem solving strategy and training of the cognitive processing capacity could be primarily accredited to WM functions. WM, the cognitive system in the prefrontal cortex of the brain, is argued to play a crucial role in developing problem solving skills because it underlies several cognitive abilities, including logical reasoning and problem solving (Vos, 2001). As described in the introduction, two aspects of WM affect the process of reasoning. (1) WM temporarily stores and manipulates information during complex cognitive activities, such as reasoning (WM capacity). (2) Prescriptive knowledge, the generic memories of effective ways to solve higher leveled cognitive problems, is stored in WM (Metacognition). We therefore categorized each function in either WM capacity or Metacognition.
**Working Memory Capacity.** The selected studies that address WM as a training intervention to specifically enhance reasoning skills in school based settings are described in Table 1.

<table>
<thead>
<tr>
<th>References</th>
<th>Sample</th>
<th>Design</th>
<th>Tasks</th>
<th>Measurements</th>
<th>Effects of training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alloway (2012)</td>
<td>15 students, 13 years middle school</td>
<td>8-week randomized control group intervention, pre-post test</td>
<td>Jungle Memory</td>
<td>Vocabulary, spelling, arithmetic</td>
<td>Significant improvement in vocabulary and math post-training, not spelling; Improvement in arithmetic</td>
</tr>
<tr>
<td>Ariës (2015)</td>
<td>63 students 15-16 years high school</td>
<td>6-week randomized passive control group intervention, pre-post test</td>
<td>Content based non adaptive n-back, odd one out</td>
<td>Reasoning in history school tests</td>
<td>No significant improvements in historical reasoning</td>
</tr>
<tr>
<td>Holmes (2009)</td>
<td>42 students 8-11 years primary education</td>
<td>7 week randomized passive control group intervention, pre-post test</td>
<td>AWMA training battery</td>
<td>Mathematical ability</td>
<td>No significant improvements in math reasoning, verbal IQ or word reading</td>
</tr>
<tr>
<td>Holmes (2014)</td>
<td>132 students 8-9 years and 9-11 years primary education</td>
<td>20-25 sessions, control group intervention, pre-post test</td>
<td>CWMT/AWMA training batteries</td>
<td>Mathematic, English</td>
<td>Significant improvements in mathematical ability and English</td>
</tr>
<tr>
<td>Karbach (2014)</td>
<td>28 students, 7-9 years primary education</td>
<td>14 sessions, randomized control group intervention, pre-post test</td>
<td>Braintwister WM training battery</td>
<td>Mathematics, reading</td>
<td>Significant improvements in reading, not math</td>
</tr>
<tr>
<td>Kroesbergen (2014)</td>
<td>51 student, 5 years preschool</td>
<td>4 week randomized passive control group intervention, pre-post test</td>
<td>AWMA battery subtests</td>
<td>Early numeracy skills, general and domain specific WM training</td>
<td>Significant improvements in domain-specific, but not domain general training groups</td>
</tr>
<tr>
<td>Loosly (2012)</td>
<td>40 students, 9-11 years, primary education</td>
<td>2-week randomized passive control group intervention, pre-post test</td>
<td>WM span task</td>
<td>Reading, nonverbal intelligence</td>
<td>Significant improvement in word reading and text reading, not in GF and pseudoword reading</td>
</tr>
<tr>
<td>Nevo (2014)</td>
<td>97 students, 8.5 mean age, primary education</td>
<td>12 sessions semi randomized control group intervention, pre-post test</td>
<td>WMP/RAP</td>
<td>Reading</td>
<td>Significant improvements on word fluency and pseudo word accuracy, not reading</td>
</tr>
<tr>
<td>Söderqvist</td>
<td>42 students, 9-</td>
<td>5-week passive</td>
<td>CWMT</td>
<td>Reading</td>
<td>Significant</td>
</tr>
</tbody>
</table>
Many different training tasks have been used in the studies with training periods that highly differ in duration. Almost all interventions have been conducted in (non-)compulsory education that precedes secondary education, ranging from preschool to middle school. Furthermore, in almost all interventions participants have been tested for verbal or mathematical reasoning.

Holmes et al. studied the improvements in mathematics and reading performance with training batteries in two studies. The first study (2009) included 22 eight to eleven year old students in primary education and tested for mathematical ability. In a 7-week training period (20-35 sessions) a randomized experimental group had to perform tasks of the AWMA training battery (see Alloway & Gathercole, 2008). Testing was conducted prior to and directly after the training period and again six months after the end of the training on standardized tests for word reading and mathematical reasoning which, according to Holmes et al., would predict academic achievement. The analysis was conducted by using a series of MANOVA’s and ANCOVA’s. The authors reported significant improvements on the posttests compared to a control group (20 participants) who completed the non-adaptive version. But, as Titz & Karbach (2014) already noted, ‘they did not include the control group into the follow-up measurements’ (p.859) which can attribute the improvements solely to maturation over the course of 6 months. Furthermore, the sample size was very small. The second study (2014) included 22 (Trial 1) and 50 (Trial 2) students in primary education. Experimental groups, specifically selected because of low-academic performance, were trained with Cogmed Working Memory Training (CWMT) batteries in 20-25 sessions (in 20 training days) and compared to a control group (50 children) which received no working memory training. Pre- and posttests contained national standard mathematic and English tests. Grade 5 and grade 6 participants showed significant improvements in math test scores comparing to control group test scores using MANOVA’s. In English test scores, only grade 6 students improved significantly compared to the control group. These results indicate that transfer between training on standard WM tasks and school test results may occur, specifically in case of low-achievers. However, the experimental group was compared to a passive control group, which might explain the differences in test results.

St. Clair-Thompson et al. (2010) used a memory strategy training (Memory booster) to explore the relationship between working memory and children’s scholastic attainment. Subgroups of 254 five to eight year old students in primary education took standardized tests in reading, arithmetic and mathematics. One class from each of the five selected schools was
assigned to the experimental condition. The other class was assigned to the control group. Experimental group pre- and posttest data (117 children) were compared with control group test results (137 children) after an 8-week intervention period, and again 5 months later. The control group received no Memory Booster training. No improvements on both posttests were observed in the standardized tests (ANOVA’s). St. Clair-Thompson et al. argue that generalizing strategies from one domain to another is not applicable to children in these age categories. The use of domain-specific content in WM training might be able to address this.

Alloway (2012) used an interactive training method based on the Jungle Memory and tested for improvements on vocabulary, spelling and arithmetic. Jungle Memory is an interactive computer program and contains three games that include spatial working memory training. The control group received low dose targeted educational support. At the beginning and end of an 8-week training period the randomized experimental (8 students) and control (7 students) groups (total of 15 participants, mean age 13) were tested. The standardized results were compared using t-tests. The training showed improvements in arithmetic abilities, but not statistically significant. However, they did find significant improvements in vocabulary and math post-training which they attribute to the interactive training regime. But only one math test was administered and the study used a small sample size which in general does not provide strong evidence of effect sizes.

Loosly, Buschkuehl, Perrig and Jaeggi (2012) used a WM span task as a primary training task in a two week randomized field study. The study investigated whether a brief intervention would result in near and far transfer effects. The span task was performed by an experimental group of 20 nine to eleven year old students in primary education and aimed to improve reading skills. Pre and posttests were conducted to compare the test results with control group data (40 children) containing standardized tests for reading abilities. One class was assigned to the experimental group condition, while selected children of from four other classes formed the control group. The children in the control group were matched with a child in the experimental group. The control group did not take part in any intervention. The intervention group received training over a period of 2 week during the first school lesson in the morning. The analyses (t-tests and MANOVA) showed that span task training improved text and word reading significantly, but not pseudoword reading. The study made use of only a small sample size and experimental group results were compared to passive control group results which might explain the differences in test results because of its novelty.

Karbach, Strobach and Schubert (2014) compared pre- and posttest data of an experimental (14 students) and a control group (14 students) who were allocated by random assignment. In 14 sessions the students in the experimental group, aged between 7 and 9 years and attending primary education, performed on a battery of adaptive Braintwister tasks. The control group received a non-adaptive low-level version of the same tasks and materials. Before and after the training they took standardized reading and mathematics tests to study the transfer of WM training to these test scores by using ANOVA on standardized scores. Significant improvements compared to pretest and control group data were found in reading, but not in math. In this regard, Titz and Karbach (2014) pointed to the mechanism that children of these age groups increasingly shift from ‘more procedural-based to more memory-based strategies’ (p.860). This forms a strong argument for differences in test scores based on new mathematical skills questions (procedural based strategies) and reading skills
(memory-based strategies). These strategy differences can also be explained by expert-novice theories which state that problem-solving skills of novices differ from experts’ problem-solving skills, which can be explained by the experts’ acquisition of schemas by which problems are solved. In this case, Sweller (1988) argues that ‘differences in memory of problem states, strategies used and categories into which problems are placed can all be explained by assuming that experts have acquired schemas which play a crucial role in the way they approach and solve problems’.

Kroesbergen, Van ‘t Noordeinde and Kolkman (2014) performed a battery of AWMA subtests in a 4-week training period (eight 30-minute training sessions) on an experimental group of five year old children in preschool. A total of 51 children were selected based on low achievements on two standardized math tests. The children were then randomly assigned to one of three conditions. The control group did not receive extra training. The intervention group received either a domain-general or a domain-specific WM training. By using pre- and posttest data in comparison to control group data Kroesbergen et al. found significant improvements in early numeracy by using domain-specific training tasks by using ANOVA. Although transfer between the training and numeracy tests has been observed, it is commonly believed that early numeracy skills do not predict secondary school reasoning skills. Furthermore, no significant improvements were found between the domain-general and the control condition on the early numeracy test. Test result differences could be explained by the extra training time which the experimental groups received. Furthermore, the small sample size makes it difficult to generalize training outcomes.

Nevo and Breznitz (2014) performed a combined training of WM and reading acceleration on students (mean age 8.5) to improve reading skills. Three study groups received a different combination of the training programs and data were compared with control group pre and posttest results. The training consisted of twelve 24-minute training sessions in three training conditions per class (each 27 children) which were compared to control group data (20 children) with no training in a series of ANOVA’s. All training programs significantly improved reading skills, such as word and pseudo-word fluency, but not reading. Nevo and Breznitz conclude that the combined use of a short WM program and a long reading acceleration program is the most effective to improve reading abilities that are related to scholastic achievement. However, the sample size is small and no characteristics of the participants were analyzed to explain test differences.

Ariës et al. (2015a) studied the effects of a content-based non-adaptive WM training on reasoning achievements in secondary school. They used two subtests of the AWMA working memory battery in a 6 week training pre- posttest design (25 minutes per week). Two classes were assigned to an experimental or control group condition According to Ariës et al. (2015a) participants in the experimental group (26 students) improved reasoning achievements compared to pretest achievements and the passive control group (16 students) in school-test results, but not significantly. This could be attributed to the non-adaptive form of the tasks. Improvements could therefore solely be attributed to more familiarity with historical content among participants of the experimental group compared to the control group.

standardized tests on math and reading were used to test for improvements. 20 students in a classroom received the Cogmed Working Memory Training (CWMT) while the control group (n = 22) received education as usual. After five weeks of training (20 training blocks) in a classroom setting under teacher guidance, the primary school students were tested. Baseline results and long term results (2 years later) were compared to passive control group results using linear regression techniques. Significant positive results were achieved on long term attainments, but not on immediate posttest results. Again, sample sizes in this study were small. Also, long term positive results of the experimental groups have not been attributed to training.

**Metacognition.** Critical thinking is related to meta-cognition, which Kuhn and Dean Jr. (2004) define as awareness and management of one’s own thought, or ‘thinking about thinking’. One can learn how to think critically by reflecting on critical thinking which makes it increasingly more effective (Kuhn, 1999; Vos, 2001; Kuhn & Dean Jr., 2004). The same qualities can be attributed to the structure of analysis concerning reasoning questions. Meta-cognitive functions are mostly executed by the WM function ‘executive control’. Consequently, WM involves executing meta-cognitive processes. Reflecting on one’s own thinking process also contributes to an increased interest in the objectives of certain activities, such as argumentation (Kuhn & Dean Jr., 2004). A meta-cognitive process can serve as a long-term learning strategy to develop reasoning skills, because the awareness of the relevance of learning activities enhances the long-term memorization and re-use of these learning activities. Kuhn and Dean Jr. (2004) define this ability as ‘internalization’. The selected studies on meta-cognitive training specified to reasoning in education are described in table 2.

<table>
<thead>
<tr>
<th>References</th>
<th>Sample</th>
<th>Design</th>
<th>Tasks</th>
<th>Type of academic achievement</th>
<th>Effects of training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ariès (2015)</td>
<td>63 students 15-17 years secondary education</td>
<td>6-week randomized control group intervention, pre-post test</td>
<td>IMPROVE</td>
<td>Reasoning in history school tests</td>
<td>Significant improvements in historical reasoning</td>
</tr>
<tr>
<td>Cardelle-Elewars (1995)</td>
<td>489 students 8-14 years primary and secondary education</td>
<td>1-year randomized control group intervention, pre-post test</td>
<td>Mayer’s model, sets of metacognitive training batteries</td>
<td>Mathematical reasoning</td>
<td>Significant improvements in mathematical reasoning independent from grade level</td>
</tr>
<tr>
<td>Kramarski (2003)</td>
<td>384 students 13 mean age secondary education</td>
<td>2-week randomized control group intervention, pre-post test</td>
<td>3 sets of self-addressed meta-cognitive questions</td>
<td>Mathematical reasoning</td>
<td>Significant improvements in cooperative settings in mathematical reasoning</td>
</tr>
<tr>
<td>Magsud (1998)</td>
<td>40 students 15-17 years</td>
<td>10-week randomized</td>
<td>Basic problem solving</td>
<td>Mathematics achievement</td>
<td>Significant improvements in</td>
</tr>
</tbody>
</table>
Four different training tasks have been used in the training studies within training periods that strongly differ in duration. Almost all interventions (excluding Cardelle-Elawar, 1995) have been conducted in secondary education. Furthermore, all interventions, except Ariës (2015a), aimed to improve mathematical reasoning.

Cardelle-Elawar (1995) used Mayer’s model and sets of metacognitive training batteries to address the effects of metacognitive instruction on mathematics achievements of low achievers. Mayer’s model includes ‘four types of processes or knowledge that are required to solve mathematics problems: translation, integration, planning and monitoring, and solution execution’ (Cardelle-Elawar, 1995. p. 82). 489 students (aged 8-14), most of Hispanic origin and low socioeconomic status, were randomly assigned to experimental and control classes and the students were given a full years’ individual training. The control group received no training. Math achievement was measured by ‘specially designed criterion-referenced tests consisting of 20 problems for each grade level’ (p.91). Tests results were analyzed by t-tests and ANOVA. Pre- and posttest results between and within groups showed significant improvements on learning outcomes in all experimental groups. This indicates the benefits of implementing metacognitive instruction in regular classrooms where a substantial share of students are low-achievers. However, because the control group received no extra training, the experimental group received 17 extra training hours.

Maqsud (1998) performed an experiment to improve mathematics achievement and attitudes towards mathematics of low achievers by using a set of basic metacognitive problem solving strategies. 40 students (aged 15-17) were randomly assigned to experimental (n=20) and control groups. The metacognitive strategies were individually trained on four mathematics topics in a 10-week training period. The control group was trained by using the conventional teaching method. Math achievements were tested by four short math achievement tests that were constructed by the teachers. By comparing pre and posttest results, Maqsud concluded that experimental group mathematics achievements improved significantly compared to pretest and control group data. However, it is not known if all classes received the same tests.

The IMPROVE method has been developed and implemented in two studies by Mevarach and Kramarski on algebra and mathematical reasoning. The first study (1997)
aimed to design IMPROVE and investigate its effects on mathematics achievement. Three classes implemented IMPROVE, while the nontreatment control group consisted of five classes. A 36-item test, covering rational numbers, identification, operation, order, laws of math operations, was administered. In a full academic year field study, they performed individual training and testing on 99 twelve and thirteen year old students of junior high school and compared the results with control group data (n=148). Experimental group students improved significantly in mathematical reasoning compared to the control groups. The second study (2003) used the IMPROVE method in both cooperative and individual training setting and compared pre- and posttest results with control groups who were instructed with worked-out examples of reasoning structures. Control group data were collected based on cooperative and individual settings. The results indicate that students who were exposed to metacognitive training outperformed both control groups. Furthermore, metacognitive cooperative training outperformed metacognitive individual training. The study indicated that lower achievers also gained more reasoning skills with the metacognitive training than with worked-out examples. Furthermore, the way groups are composed and group interaction is structured is important for the training effects, i.e. effects were largest for small groups of 4 participants consisting of an overachiever, middle achiever, underachiever.

Kramarski and Mevarach (2003) studied the effects of 4 instructional methods on mathematical reasoning and meta-cognitive knowledge: cooperative learning including meta-cognitive training (COOP+META); individual learning including meta-cognitive training (IND+META); cooperative learning excluding meta-cognitive training (COOP); individual learning excluding meta-cognitive training (IND). A group of 384 students, average age 13, was randomly selected from 4 secondary schools, which in turn were randomly selected from 15 secondary schools in a single district in Israel. The schools were comparable in size and social economic status. Twelve certified female teachers with a minimum of 5 years of experience taught during the training period. Students were trained in mathematical problem-solving in all classes for 5 times a week during a period of 2 weeks. The meta-cognitive training was based on 3 sets of meta-cognitive questions: comprehensible questions; strategic questions; connection questions. All students were tested before, during and after the training period. COOP+META used small heterogeneous groups that had to respond to the 3 sets of meta-cognitive questions. IND-META used the same sets of meta-cognitive questions, but students had to respond to them individually. COOP used small heterogeneous groups that didn’t respond to math problems by using the meta-cognitive questions. IND used individual response without meta-cognitive questions. Kramarski and Mevarach found significant differences in mathematical reasoning. COOP+META scored better on posttests than IND+META which, in turn, scored better on posttests than COOP and IND. There were no significant differences between COOP and IND in achievements on posttests. Significant differences were found in fluency in reasoning and flexible reasoning. Meta-cognitive groups achieved significantly better than non-meta-cognitive groups on a transfer task. But there were no significant differences between the meta-cognitive groups and the non-meta-cognitive groups.

Ariës et al. (2015a) studied the effects of a 6-week metacognitive history training on reasoning achievements in secondary school history tests. Participants were trained in answering reasoning questions of old history exams by using the IMPROVE method in a 6
week training pre- posttest design (25 minutes per week). The IMPROVE method consists of a set of metacognitive questions which participants had to address in a cooperative form. Pre- and posttest data (school tests) were compared to a passive (traditional schooling) and an active (WM capacity training) control group. Participants in the experimental group significantly improved reasoning achievements compared to pretest achievements and both control groups after three weeks of training.

Discussion and conclusion

This review contributes to the literature on enhancing problem solving skills by focusing on whether a combined brain-based training method of WM capacity and metacognition is effective and can be implemented in primary and secondary education for students to develop problem solving skills. In the discussion and conclusion section, we therefore specifically focus on outcomes in problem-solving.

The WM-capacity learning strategy implicates the enhancement of WM which is argued to correlate with improved reasoning abilities. The research reviewed above shows that training tasks and training intensity strongly differ between studies. Participants’ ages ranged from 5 to 16 years attending primary or secondary education from preschool to secondary school. Almost all studies included training of mathematics abilities and/or reading and all but one study (Ariës et al. 2015a) included general (non-content-based) training tasks. 50% of the interventions in the studies we included in this review resulted in significant improvements in reasoning achievement, regardless of the type and intensity of training or age. Significant improvements occurred more often in reading compared to math tests.

In the introduction we pointed out that the transfer to reasoning achievement is likely to be more effective when participants are trained by using content-based interventions due to near-transfer effects. Near transfer effects occur when the content of the training batteries is similar, but not identical, to test content. In this regard, it seems rather peculiar that the Ariës et al. (2015a) study is one of the few studies that didn’t result in significant improvements in reasoning achievements. However, this can be attributed to the absence of an adaptive element in the training tasks, which is considered a core element in WM training, and to the cognitive levels of the participants in this study who were students in preparatory academic education (the highest track in Dutch secondary education) and therefore already had above average WM levels. The latter argument is in line with the compensation account theory of which Titz and Karbach (2014) state that ‘high-performing individuals will benefit less from cognitive interventions, because they are already functioning at the optimal level, which leaves less room for improvement’ (p. 861). One study on reading comprehension used content based training batteries in which near transfer effects could occur. In the Nevo et al. (2014) study, near transfer effects occurred in vocabulary, but not in reading. Two studies concerning mathematics used content based training batteries in which near transfer effects could occur. Alloway et al. (2012) found positive results in mathematical reasoning, but had a small sample size. In the Kroesbergen et al. (2014) study, content based training resulted in significant improvements in early numeracy.

All other studies used general training tasks, which could provide evidence of far transfer of standardized training to reasoning achievements in school tests. Far transfer effects occur when the training content is not similar to the tests, yet improvements are found
that can be attributed to training. The results that we found in studies that contained general training batteries prove to be more ambiguous. With regard to reading four studies found positive relationships with regard to vocabulary (Holmes et al., 2014; Loosly, 2012), spelling (Loosly, 2012) and reading comprehension (Holmes, 2014; Karbach, 2014; Loosly, 2012; Söderqvist 2015). However, these results were compared with passive control groups. Two studies (Holmes, 2009; StClair-Thompson, 2010) found no results in reading comprehension. Alloway et al. (2012) did find a significant improvement in vocabulary, but not in spelling. The ambiguous pattern of results in the Alloway et al. (2012) study, however, can be most likely attributed to the small sample size. With regard to mathematics, far transfer effects between general training and tests are less likely to occur. This could be attributed to the specifics of reasoning in mathematics in which students reason at an abstract level, which can not be compared to reading comprehension or reasoning in social sciences in which students mainly deal with concrete topics. Two studies (Holmes, 2014; Söderqvist, 2015) found positive relations with math reasoning. Three studies (Holmes, 2009; Karbach, 2014, StClair-Thompson, 2010) did not found significant improvements on math reasoning or early numeracy (Kroesbergen, 2014).

Most participants did attend primary education and not secondary education during the intervention. Therefore, further research is needed to find out whether the existing training designs can be copied to secondary education settings in which course specific heuristics and meta-concepts are more important than in primary education. The articles included in this review by Alloway et. al. (2012), Holmes et. al. (2009), Holmes & Gathercole (2014), and Kroesbergen et. al. (2014) all included samples of children with low performance, which could suggest that the results can not be generalized to all children. However, we could not determine if the results of these studies differed from those using samples of children with normal learning behavior.

Reading comprehension skills play a significant role in social sciences, e.g. for history in interpreting and synthesizing historical sources to build an interpretative case, in which the historical sources frequently contain written documents or, to a lesser extent, data in the form of tables and graphs. In this regard, the reviewed studies could be of value in enhancing reasoning skills in social sciences courses.

The meta-cognitive learning strategy trains the management and control over one’s own thinking and internalization. The awareness of the relevance of learning activities results in long term memory storage by which learning activities can constantly be re-used. This is specifically applicable to achievements in general reasoning, but also applies to fluent reasoning, flexible reasoning and the transfer of reasoning. The reviewed articles show that students achieve significantly better on reasoning tests when meta-cognitive skills are trained compared to students who have had no meta-cognitive training. The findings that are reviewed in the previous paragraphs show that training tasks and training intensity strongly differed. The participants’ ages ranged from 8 to 17 years attending compulsory education from primary to secondary school. All studies but one (Ariës et al., 2015a) included training on mathematics abilities and used general (non-content-based) training tasks. All interventions reviewed in this paper resulted in significant improvements in reasoning achievements, regardless of the type and intensity of training or age.
Two instructional designs on meta-cognition have been reviewed: the method in which small heterogeneous groups work together in a meta-cognitive process and the method in which meta-cognition is trained individually. Heterogeneous groups that are trained in meta-cognition turn out to achieve better compared to individual training methods on general reasoning achievements, fluent reasoning and flexible reasoning. But this significant difference cannot be applied to the transfer of reasoning between different school courses without more research on this specific topic.

The review further suggests that five issues need to be addressed in meta-cognitive training methods for problem solving structures to be internalized in long term memory. (1) Participants need to solve problems by using a training method which includes both self-reflection and reflection proposed by group members. (2) As stated in the introduction, evidence suggests that problem solving structures need to be trained by using content based knowledge on reasoning questions (e.g. Van der Sluis et al., 2007). (3) Problem solving skills can best be internalized by training in heterogeneous groups, containing underachievers, overachievers, and average achievers in problem solving skills. The group is responsible for reflecting on a participants’ problem solving process. (4) Skills are best internalized when a participant follows a step-by-step plan and verbally presents the reasoning process to the other group members. (5) Training needs to be consistently repeated. When implementing the IMPROVE method, Mayer’s model or subsets of metacognitive training batteries in a cooperative form, addressing content based reasoning questions, and repeated (weekly) training, then all criteria are accounted for.

In line with Leinhardt et al. (1994) we suggest that content based knowledge may be crucial to acquire problem skills because of the alternative views, detailed factual knowledge, broader frames of reference and other concepts that are frequently used in compulsory education courses. This contradicts statements suggesting that transfer will occur to content based problem solving skills when students are only trained on standardized tests. In fact, studies indicate that the implementation of subject-oriented content and processes is an important element in enhancing the effectiveness of EF and WM training models (Vos, 2001; Van der Sluis et al., 2007; Lee et al., 2009). More research has to be conducted to link standardized working memory tests to content based knowledge in which students are trained in. Furthermore, the implications of content-based training methods, contrary to standardized training methods, should be studied.

The students included in the studies include primary school and secondary school aged children. Few WM capacity studies have been conducted on a considerable part of the target group of the review: adolescents. Most of the research on this matter has been conducted on children and adults. This means that no firm conclusions can be drawn about training of cognitive reasoning abilities for adolescents. However, because research on children and adults proved to have similar results, a tentative conclusion is that similar training methods can enhance adolescents’ reasoning abilities. More research on the adolescents’ brain in relation to brain based training methods needs to be conducted to provide more insight on this matter.

Another practical problem lies in the development of the adolescent brain, specifically the working level. The training model has to account for differences in working levels
between different students and the use of a longer training period implies that the model should grow as the brain matures. The question is: how? Should a model be developed per level or class or students’ age? How long should students be trained in order to become effective? When is training not effective anymore? Can individual training practically be integrated in curricula? These are practical and financial issues that could be obstacles to implement these training methods into educational practice.

Practical restrictions to implement general working memory training in school curricula make it difficult to integrate general and content-based cognitive training methods in school practice. For instance, to what extent are students motivated when being trained by a general method to enhance content based reasoning? Can teachers be convinced that they should train general reasoning methods to students instead of using only content-based interventions? Therefore, research has to be conducted on how to implement content based problem solving training into working memory training models. With regard to concentration, training methods require that students focus their attention on an internal representation. Attention has two components: (1) an inclusionary component consisting in the content in focus, and (2) an exclusionary component consisting on the inhibitory control of interference. Since WM is demonstrably enhanced by concentration and weakened by distractibility, underachieving could easily been caused by distractions (Ariës et al., 2015b). Therefore, one would consider minimizing distractions as a requirement of an effective methodology. Furthermore, many studies lack a thorough critical analysis of the mechanisms by which WM causes reasoning achievements. For instance, training advantages may be caused by automated skills, other than WM, that reduced the load on WM. Or training might improve concentration or improve the motivation of the training group. These possibilities are rarely considered but undermine the effects of WM training on WM capacity enhancements. Future research has to be conducted on the added value of course-specific training of working memory functions compared to domain-general training to enhance problem solving skills.

Acknowledgements

We gratefully acknowledge the contributions of Joaquín Fuster, UCLA Semel Institute for Neuroscience & Human Behavior, and Allyson P. Mackey, MIT McGovern Institute for Brain Research, who supported and gave extensive feedback on the description of the mechanisms and implications of the training methods. This work was supported by the Netherlands Organisation for Scientific Research (NWO) under the Doctoral Grant for Teachers.
References


Goldberg, E. (2010). *Het sturende brein: Onze hersenen in een complexe wereld [The
new executive brain: Frontal lobes in a complex world]. Amsterdam, Wereldbibliotheek.


