An evidence-based approach to working-memory based training in secondary education to improve reasoning test achievements

Summary

Teachers regularly observe that adolescent students have insufficient reasoning skills to properly answer reasoning test questions. The importance of acquiring reasoning skills in secondary education becomes clear in humanities school-tests which increasingly contain reasoning questions. Policymakers and schools are searching for new didactical approaches to improve adolescents' reasoning skills. One potentially effective strategy is developing learning strategies based on cognitive and brain science, popularized as 'brain-based learning' (Crone et al., 2016). Standardized training batteries have been developed that are theorized to improve reasoning performance while they focus on areas of the brain that deal with reasoning, such as working memory (WM). But these batteries are not yet being implemented in school curricula, and school-based WM training to improve reasoning is still limited and recent. Evidence of improved reasoning in education, caused by WM training, therefore remains poor. This thesis contributes to our knowledge and insight into effective brain-based strategies that improve reasoning skills in secondary education humanities curricula. We evaluate the effectiveness of a combined WM-capacity- and reasoning-strategy training which enables students to acquire reasoning skills to improve their academic achievement. This could provide the necessary tools for both teachers and students to improve school-test results in humanities courses.

Humanities courses in secondary education in the Netherlands include history and social studies, among others. Tests usually consist of deductive reasoning problems. The required problem-solving skills are generally comparable for all humanities courses. They consist of general problem-solving principles, such as analysis, hypothesis, generation, and interpretation, and of domain-specific characteristics that are specific to the particular courses, such as domain-specific knowledge and the meta-concepts of the domain. Each teacher teaches the required knowledge and skills to the students, primarily by using direct instruction, course

book questions and (un)official school tests in order to train problem-solving skills. In this regard, the student should be taught to reorganize and manipulate retained information, which are seen as skills that are highly important to solve reasoning questions. Developing a problem-solving strategy and training in cognitive processing capacity can be primarily accredited to WM functions (Fuster, 2003). WM underlies several cognitive abilities, including logical reasoning and problem-solving. But 'the cortical hubs of the network [...] are slow to mature' and sensitive to genetic and environmental factors, which could explain individual differences in reasoning test achievements.

We hypothesize that WM training is only effective in educational settings when accompanied by reasoning strategy training that is subject-based in the humanities. In Chapters 2 and 3 we present review studies that examine the possible effects of a single WM-capacity training and a combined WM-capacity- and reasoning-strategy training. These two chapters set out to address the following questions:

1) Does n-back training improve school-based reasoning achievements?;
2) Can WM capacity + metacognitive training serve as an effective instrument to enhance school-based reasoning achievements?

Chapters 4 and 5 focus on the effect of the combined WM-capacity- and reasoning-strategy training in history and in humanities education. In these chapters, we answer the following questions:

3) Does WM capacity + metacognition training, based on subject-matter knowledge in history education, have an effect on students' achievements in tests for which reasoning abilities are required?;
4) Does WM capacity and/or metacognitive training, when independently trained, have an effect on students' achievements in history tests for which reasoning abilities are required?;

5) Does the combined training have an effect on students' reasoning achievements in school-tests in other humanities courses (i.e. social science)?

Chapter 6 includes the combined intervention in which motivational factors are included to investigate the added value of motivation during the training period. The following question is addressed:

6) Does the combined training affect reasoning achievements in school-test results when motivational factors are added to the training?

Chapter 2 of this dissertation reviews the effects of one specific WM training task, the n-back, on deductive reasoning improvements. The n-back task involves the serial presentation of a stimulus, spaced several seconds apart. The participant must decide whether the current stimulus matches the one displayed *n* trials ago, where *n* is a variable number that can be adjusted up or down to, respectively, increase or decrease cognitive load. In the context of WM training, efforts have focused on flexibly adapting the task difficulty in accordance with the participants' fluctuating performance level by increasing and decreasing the level of *n*. The idea is to keep the participant's WM system perpetually engaged at its limit, thereby stimulating an increase in WM function.

Specifically n-back training, which is a commonly used task in WM task-transfer training studies, could lead to positive effects on general fluid intelligence (*Gf)* (Au et al., 2015; Jaeggi et al., 2008; Beatty and Vartanian, 2015). Ergo, n-back training could theoretically improve deductive reasoning via its effect on *Gf.* N-back requires storage and updating of information in WM and 'interference resolution'. Therefore, more than other WM-training tasks, n-back training could positively affect fluid intelligence (Au et al., 2015). N-back training might stimulate the WM function and, as a result, improve reasoning. In this chapter, we review the literature on the effects of the n-back on deductive reasoning improvements. To enlarge on far-transfer results to reasoning, we

included studies on school-based and non-school-based interventions in our review.

The main characteristic is that interventions must: include some form of n-back and have far-transfer measures to (school-based) reasoning; use a control group; and include a deductive reasoning outcome measure. The literature review shows that all interventions have used a pretest-intervention-posttest design, in which objective academic measurements are used. All but one intervention (Jaeggi et al., 2011) used randomized experimental and control groups. The average sample size was 67 participants. All interventions used identical pretest and posttests to measure for far transfer effects. We have found five significant and eight non-significant results concerning far-transfer effects to reasoning (*Mdn d* = 0.10*, range* -0.35 − 2.20*).*

Four interventions were school-based. This means that the training was implemented in the school curricula. Two studies have found significant results and two studies did not (*Mdn d* = 1.04*, range* 0.66 − 2.20). In conclusion, no far transfer occurs when solely n-back is trained. When n-back training is combined with other training tasks, then the results appear to be ambiguous with respect to far-transfer to reasoning. Eight non-school-based interventions were included in the review. Five out of eight interventions observed no far-transfer to reasoning posttests (*Mdn d*= 0.16*, range* -0.32 − 0.80). Furthermore, significant improvements in posttest results did not occur when compared with the results of an active control group. Many studies lack a critical analysis of the mechanisms that cause reasoning achievements through WM or n-back training, such as skill internalization, other than WM (see also: Lilienthal et al., 2013, p.135). In school-based settings, when WM training is combined with reasoning strategy training, significant improvements do occur in reasoning-test results. One might question the role of n-back in this transfer to reasoning, and whether reasoning strategy training is primarily responsible for transfer results. In this regard, it might be the case that WM (improvements) functions as a subsystem that supports improvements in reasoning through reasoning strategy training. Five of

seven studies in the young adult group used samples of university students. One might consider that this is a rather homogeneous group who already had above average WM levels, a situation which has been theorized to limit WM training benefits. The possible benefits of a combined WM- and reasoning-strategy training to reasoning test-achievements in primary and secondary education is addressed in Chapter 3.

Chapter 3 reviews whether a combined training of WM capacity and reasoning strategy could benefit school-based reasoning skills, while two aspects of WM are theorized to 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 important for problem-solving. (2) Prescriptive knowledge, the generic memories of effective ways to solve higher-level cognitive problems, is stored in WM (Gazzaniga et al., 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 WM 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. Meta-cognitive training causes WM to form blueprints and to plan for rational analysis and analytical methods (Gold et al., 1996; Gazzaniga et al., 2009; Goldberg, 2010).

We have found a total of 16 articles (10 articles concerning WM-capacity training, and 6 articles concerning metacognitive training)

eligible for inclusion in the systematic review. Many different WM-capacity training tasks were used in the studies with training periods that highly differed in duration. Almost all interventions were conducted in (non-)compulsory education, ranging from preschool to middle school. Furthermore, in almost all interventions, the participants were tested for verbal or mathematical reasoning. 50% of the interventions 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 with math tests. 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. Two studies on 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. The results that we found in studies that contained general training batteries proved to be more ambiguous. 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 with reading comprehension or reasoning in the humanities, in which students mainly deal with concrete topics.

As for the metacognitive training, four different training tasks were used in the training studies within training periods that greatly differ in duration. Almost all interventions were conducted in secondary education. Two instructional designs on meta-cognition were 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 with individual training

methods on 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 in order 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 (Chapter 2), evidence suggests that problem-solving structures need to be trained by using content-based knowledge on reasoning questions. (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.

We conclude that both instructional designs could improve reasoning skills. In line with Leinhardt et al. (1994), we suggest that content-based knowledge may be crucial to acquire problem-solving skills because of the alternative views, detailed factual knowledge, broader frames of reference, and other concepts that are frequently used in compulsory education courses. Most of the research on this matter has been conducted on children and adults. This means that no firm conclusions can be drawn about the training of cognitive reasoning abilities for adolescents. Practical restrictions to implement general WM training in school curricula make it difficult to integrate general and content-based cognitive training methods in school practice. The question how to implement the combined training and its effects on improving reasoning-test achievements is addressed in Chapters 4, 5 and 6.

Chapter 4 evaluates two content-based interventions in the history courses of higher general secondary education (HAVO) and pre-university

secondary education (VWO). HAVO students (total N=92) received the combined WM-capacity- and reasoning-strategy training battery in order to analyze its effectiveness in their reasoning achievements in school-tests. WM capacity included n-back and Odd One Out tasks, while reasoning strategy training is performed by using the IMPROVE-method. The IMPROVE-method is a format to develop and structure the reasoning process (Mevarech and Kramarski, 2003). The results are compared with a control group who received a conservative training. WM capacity, reasoning strategies, and reasoning achievements in the school tests of the experimental group improve significantly after 3 weeks of training. The gain in reasoning abilities is also durable and remains significant up to 16 weeks after the training program is completed.

In VWO (total N=63), WM capacity and reasoning strategies are independently trained and compared with control group results to analyze the improvements of both training groups independently. Training of WM capacity did not improve achievements in reasoning tests significantly, but training of reasoning strategies did. The results of both experiments cannot be compared because of cognitive group differences between Experiments 1 and 2 (in terms of educational tracks). These results demonstrate that brain based learning strategies to improve reasoning abilities can easily be integrated in history curricula. Furthermore, Chapter 4 shows that a 6-week training period can significantly improve reasoning abilities of students in secondary education.

Chapter 5 includes the combined intervention in HAVO in social studies courses. The main aim of this intervention is to analyze whether the training duplicates the results of the intervention described in Chapter 4. This is an important question which can give insight into the extent to which the training is applicable in all the arts courses, as well as possible transfer effects which can be investigated in the future, as both history and social studies are courses taught in the humanities curricula.

In three secondary schools in the Netherlands we conducted an experiment that is based on WM training tasks. The aim of the training intervention was to: (1) investigate the effect of the complete training on achievement in reasoning tests; (2) investigate its durability after completion of the training; and (3) investigate the effects of age, gender, pretest reasoning-achievement scores, and WM pretest scores on posttest achievements. The experimental groups' test results (N=39) were compared with the data of the control groups (N=42) that were trained by using 'low-dose' WM tasks, such as evaluating reasoning questions using whole-class teaching methods.

WM-capacity and reasoning achievements improved significantly in the experimental groups after 4 training weeks compared with the control groups. The gain in reasoning abilities was demonstrated in both experimental subgroups and remained significant 8 weeks after the end of training, while the control groups' results did not improve. The experimental groups' improvements are in agreement with conclusions of Chapter 4, in which similar groups underwent the combined training in history curricula. The results of the regression model indicate that, for every unit increase of pretest achievement scores, the score of the reasoning achievements posttest can be predicted to increase by .29 units. Students with a low pretest reasoning-achievement score typically improve less in the posttest compared with high pretest score students. This has important implications for the effectiveness of the training for underachieving students. While it is possible that the training tasks could be going too fast for these students, the absence of an improvement in test scores could also be caused by behavioral characteristics, such as lack of motivation or concentration. With regard to the latter, the training method requires that students focus their attention on an internal representation. Although motivation and concentration were not analyzed, minimizing distractions may be a requirement for an effective intervention. Furthermore, enhancing and training attention to improve deductive reasoning could also be regarded as an important outcome measure of the training, while it is also regarded as a general cognitive

function within which WM is subsumed. To investigate the role of motivation in the training, we added motivational components in a second experiment in history courses in Chapter 6.

Chapter 6 contains an analysis the combined intervention to which motivational factors are added in order to investigate the results of the less-motivated students. In this experiment, we compared the school test results of both groups who had received the combined training, where one group was actively being motivated by the teacher by use of verbal praise, goal-setting, and feedback, which are known to improve attention. Both the HAVO experimental (N=50) and the control (N=58) groups received the training in history courses. The main objective was to analyze whether the addition of motivational components can improve test results even more compared with the control groups.

Group motivation levels are measured by comparing each group's pretest and posttest scores regarding a motivation questionnaire. We observed a significant decline of motivation levels in the control group. But neither were the motivation levels in the experimental group significantly affected. These results indicate that training on WM and reasoning strategies might require huge cognitive loads that can easily affect motivation levels. The results indicate that the motivation incentives that were used during the training in the experimental group were shown to be effective in maintaining motivation. However, the higher levels of motivation do not account for significant improved reasoning test scores. Our conclusion is that motivation levels do not predict effects of WM- and reasoning-strategy training. This allows teachers to concentrate on the learning process of the students rather than investing time in keeping the students motivated.

In Chapter 7, the findings of the review and the empirical studies lead to the formulation of four key conclusions.

(1) There is no convincing evidence that (content-based) WM capacity training alone improves adolescents' reasoning skills in education;

(2) Combined content-based WM-capacity- and reasoning-strategy training appears to be an effective instrument to enhance school-based reasoning achievement;

(3) Reasoning achievements in school-test results do not seem to be affected by motivation levels when students use the combined training method;

(4) The combined WM-capacity- and reasoning-strategy training does not appear to discriminate between low or high achieving students in terms of improvements on reasoning tests.

A number of methodological issues and policy implications arise from our findings. With regard to the methodological issues, we propose that non-school-based WM training can not be implemented one-on-one in school curricula in order to train reasoning skills, because of the absence of other aspects, such as knowledge of contents and course-specific skill acquisition, which are regarded as being important contributors to school-based reasoning-test results. We argue that a general non-school-based WM training does not provide the necessary transfer effects to improve the children's reasoning skills to enable them to improve significantly in school-based reasoning tests. Consequently, standard WM training tasks should be tailored to suit the demands of school-based reasoning. Another concern addressed in this PhD-thesis relates to the development of an effective training method to improve reasoning skills and reasoning achievements for students of varying backgrounds and baseline levels. Our findings do not imply that specific groups of students (based on low- or high-achieving baseline levels, gender, socioeconomic status, general cognitive ability or motivation levels) particularly benefit from the training compared to other groups of students. We conclude that a one-size-fits-all training is not likely to include all students in the possible benefits of the training.

With regard to the policy implications, we conclude that our training method is a much more effective instrument for students to acquire the necessary reasoning skills needed to adequately respond to a reasoning question (*d = 1.04-2.20)*, compared with frontal instruction designs. We propose that the explanation for this difference in outcomes could be that, besides improvements in WM levels, the current training engages all students actively in solving reasoning questions and in WM-capacity training. Reasoning skills are trained more effectively by the use of active didactic approaches. We also argue that the training in general could be implemented in high school curricula to effectively improve the students' reasoning skills in the humanities. However, the training does not provide a clear-cut instrument to target specific groups who need to be trained by having personal learning designs. We also want to emphasize the importance of integrating a standard form of reasoning training in secondary education to show students that the same reasoning structures can be applied in different courses. By creating awareness in this matter, schools can offer more efficient and effective learning strategies to students, and invest in interdisciplinary education concerning the acquisition of reasoning skills. In this case, the findings imply that the training, when focused on course-based content, can be implemented in all courses in the humanities, such as history and social studies, to improve reasoning test achievements.