

KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Onderzoekseenheid Gezins- en Orthopedagogiek

**Associations between inhibition and arithmetic
fact retrieval**

Master's thesis submitted for the
degree of Master of Science in de
pedagogische wetenschappen by
Elien Bellon

Supervisor: Prof. Dr. Bert De Smedt

2015

Summary

Mathematical skills are crucial abilities in modern western society, so it is important to uncover the cognitive processes underlying children's achievement in mathematics. Why are some children very good in doing calculations and do others experience life-long difficulties with this basic competency? Most research in the last decade has identified the ability to process numerical magnitudes as a core determinant of these individual differences in mathematics achievement. However, this exclusive focus on one core domain-specific cognitive factor for explaining mathematical (dis)abilities has been recently criticized because it ignores the involvement of other critical cognitive functions and processes that might play a role in the development of mathematical skills.

This study explores the role of inhibition in arithmetic fact retrieval. Specifically, we examined if inhibition is associated with arithmetic fact retrieval, and we investigated the unique role of inhibition in determining individual differences in arithmetic fact retrieval, above numerical magnitude processing. Inhibition might play a role in fact retrieval, because when we retrieve arithmetic facts from memory, incorrect (related) answers need to be inhibited. On the other hand, we focussed on arithmetic fact retrieval because this is a major building block for further mathematical development, and it is a characteristic deficit in children with mathematical disabilities. We used a correlational analysis, regression analysis, and cluster analysis to examine the association between measures of inhibition and arithmetic fact retrieval.

Our study failed to observe a significant association between inhibition and arithmetic fact retrieval. Consequently, our results did not reveal a unique contribution of inhibition to arithmetic fact retrieval above numerical magnitude processing. Symbolic numerical magnitude processing was significantly associated with arithmetic fact retrieval. Future studies should further analyse the association between arithmetic fact retrieval and inhibition, for example by considering strategy use when solving arithmetic problems, by providing a more careful characterization of the inhibitory processes involved in arithmetic fact retrieval through the use of tasks that tap into different components of inhibition, by examining these associations in older children and in atypically developing groups (e.g., ADHD), and by investigating this association at the neural level.

Content

Abstract	1
Introduction	2
Numerical magnitude processing	4
Inhibition	5
The current study	7
Method	9
Participants	9
Materials	9
Numerical magnitude processing.	10
Inhibition.	10
Stroop tasks.	10
BRIEF.	11
Arithmetic fact retrieval.	11
Verification tasks.	11
Tempo Test Arithmetic.	12
General mathematics achievement.	12
Reading ability.	13
Control measures.	13
Intellectual ability.	13
Motor reaction time.	13
Procedure	14
Results	14
Descriptive analyses	15
Correlational analyses	16
Regression analyses	17
Cluster analyses	19
Discussion	23
Limitations and future directions	27

References	31
Appendix	47
1. List of problems administered in the arithmetic verification tasks	48
2. Missing data	50
3. Supplementary results	51
Accuracy data.	51
Sex differences.	52
Assumptions regression analysis.	53
Assumptions ANOVA.	60
Assumptions ANCOVA.	61

Abstract

Although it has been proposed that inhibition is related to individual differences in mathematical achievement, it is not clear how it is related to specific aspects of mathematical skills, such as arithmetic fact retrieval. The present study therefore investigated the association between inhibition and arithmetic fact retrieval and further examined the unique role of inhibition in individual differences in arithmetic fact retrieval, above numerical magnitude processing. We administered measures of inhibition (i.e., stroop tasks and a teacher questionnaire) and arithmetic fact retrieval (i.e., a production and two verification tasks) in 95 typically developing third graders. We used a correlation, regression, and cluster analysis. This study failed to observe a significant association between inhibition and arithmetic fact retrieval. Consequently, our results did not reveal a unique contribution of inhibition to arithmetic fact retrieval above numerical magnitude processing. However, the significant unique contribution of symbolic numerical magnitude processing to arithmetic fact retrieval was confirmed. Several directions for further research are discussed.

Keywords: mathematical competencies, individual differences, arithmetic fact retrieval, inhibition, numerical magnitude processing, third grade.

Introduction

There are large individual differences in the way children acquire mathematical competencies (e.g., Dowker, 2005). Mathematical skills are crucial abilities in modern western society (e.g., Ancker & Kaufman, 2007; Finnie & Meng, 2001). Our daily life is permeated by numbers and one's success in dealing with numbers and quantities is related to job prospects, income and quality of life (e.g., poor mathematical competencies lead to greater risk of underemployment) (Parsons & Bynner, 2005). Early mathematical skills predict later adult socioeconomic status (Ritchie & Bates, 2013). As a result, it is important to understand these individual differences and the cognitive processes underlying children's achievement in mathematics. This understanding can contribute to designing scientifically validated diagnostic tests and remediation programs for children at risk for or with difficulties in mathematical achievement.

The successful learning and performance of mathematics relies on a range of social (e.g., Byrnes & Wasik, 2009), educational (e.g., Opdenakker & Van Damme, 2007) and individual factors (Cragg & Gilmore, 2014). Individual factors contributing to differences in mathematics achievement include non-cognitive factors, such as attitudes (Ma, 1999) or motivation (Steinmayr & Spinath, 2009) as well as cognitive factors, such as numerical knowledge (De Smedt, Noël, Gilmore, & Ansari, 2013; Jordan, Glutting, & Ramineni, 2010), working memory and executive function (Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013), language (Donlan, Cowan, Newton, & Lloyd, 2007) and intellectual ability (Mayes, Calhoun, Bixler, & Zimmerman, 2009).

In this study, we focussed on the cognitive determinants of these individual differences in mathematical skills. There are two dominant ways to study these cognitive determinants, i.e. a domain-specific and a domain-general approach (e.g., Fias, Menon, & Szucs, 2013). Domain-specific approaches investigate the role of number-specific processes, such as the representation of numerical magnitudes in individual differences in mathematics achievement (e.g., De Smedt et al., 2013). Domain-general approaches focus on the influence of non-numerical cognitive skills, such as phonological skills (Hecht, Torgesen, Wagner, & Rashotte, 2001), inhibition (Allan, Hume, Allan, Farrington, & Lonigan, 2014; Bull & Scerif, 2001; Gilmore et al., 2013), working memory (De Smedt, Verschaffel, &

Ghesquière, 2009; Raghobar, Barnes, & Hecht, 2010), retrieval from long-term memory (Dowker, 2005) and visuo-spatial processing (Geary, 1993) on mathematical performance.

Recently research has mainly focussed on domain-specific factors, thereby ignoring more domain-general factors, and only a few studies focussed on both domain-specific and domain-general factors. Also, most research mainly investigated mathematics performance with broad general standardized achievement tests, which typically assess a wide variety of mathematical skills (e.g., arithmetic, problem solving, geometry) and yield a total score that reflects performance averaged across various mathematical domains. This does not allow to carefully pinpoint the associations with different mathematical skills. The current study extends abovementioned research by examining the association between one specific mathematical skill (i.e., arithmetic fact retrieval) and a specific domain-general cognitive skill (i.e., inhibition), taking into account the influence of a domain-specific skill (i.e., numerical magnitude processing). Such research is needed, in order to more carefully pinpoint associations between cognitive and mathematical skills.

There are several important reasons for focussing on arithmetic fact retrieval. Firstly, mathematics is a cumulative skill. Being skilful at single-digit arithmetic and fact retrieval is a major building block for further mathematical development (e.g., Campbell & Xue, 2001; Kilpatrick, Swafford, & Findell, 2001; Koponen, Salmi, Eklund, & Aro, 2013; Vanbinst, Ceulemans, Ghesquière, & De Smedt, 2015). Secondly, children with mathematical learning difficulties consistently show deficits in arithmetic, namely in executing procedures and in fact retrieval (e.g., Barouillet & Lépine, 2005; Geary, 2004; Geary, Hoard, & Bailey, 2012; Jordan, Hanich & Kaplan, 2003; Landerl, Bevan, & Butterworth, 2004).

In the remainder of the Introduction, we first review the available evidence on the association between numerical magnitude processing and arithmetic fact retrieval. Next, we review the studies that have examined the association between inhibition and individual differences in mathematical skills. Finally, we present the specific aims of our study.

Numerical magnitude processing

Numerical magnitude processing, or *people's elementary intuitions about quantity and their ability to understand the meaning of symbolic numbers* (Vanbinst, Ceulemans, et al., 2015, p. 30), is one major determinant of individual differences in mathematical performance on which research over the last five years has intensively focussed (De Smedt et al., 2013, for a review). Various studies have shown that performance on numerical magnitude comparison tasks – often used as a measure of numerical magnitude processing – is correlated with (e.g., Durand, Hulme, Larkin, & Snowling, 2005; Holloway & Ansari, 2009; Sasanguie, Gobel, Moll, Smets, & Reynvoet, 2013) and even predicts (De Smedt, Verschaffel, et al., 2009; Halberda, Mazocco, & Feigenson, 2008; Vanbinst, Ghesquière, & De Smedt, 2015) individual differences in mathematical achievement. This association is observed on both symbolic comparison tasks, consisting of Arabic digits stimuli (De Smedt et al., 2013; De Smedt, Verschaffel, et al., 2009; Durand et al., 2005; Holloway & Ansari, 2009; Sasanguie et al., 2013) and non-symbolic comparison tasks, consisting of dot arrays (Halberda et al., 2008; Mundy & Gilmore, 2009). Also, children with mathematical difficulties have particular impairments in understanding and processing numerical magnitudes (De Smedt et al., 2013, for a review).

Numerical magnitude processing has also been associated with specific mathematical skills, such as arithmetic fact retrieval (De Smedt & Gilmore, 2011; Geary, 2010; Robinson, Menchetti, & Torgesen, 2002; Vanbinst, Ceulemans, et al., 2015; Vanbinst, Ghesquière, & De Smedt, 2012; Vanbinst, Ghesquière, et al., 2015). Vanbinst and colleagues (2012) found that children who had better symbolic magnitude representations retrieved more facts from their memory. Moreover, Vanbinst, Ghesquière and De Smedt (2015) found that children's numerical magnitude processing skills even predict individual differences in arithmetic. On the other hand, impairments in numerical magnitude representation are directly related to poor performance in single-digit arithmetic (De Smedt, Reynvoet, et al., 2009). There is also neural evidence for the importance of numerical magnitude processing for higher level mathematical tasks such as arithmetic. For example, cognitive neuroimaging studies have shown that the intraparietal sulcus – dedicated to the processing of magnitudes (Ansari, 2008, for

a review) – appears to be consistently active during arithmetic tasks (Menon, 2015, for a review).

However, this recent exclusive focus on one domain-specific cognitive factor for explaining individual differences in mathematical performance has been recently criticized (e.g., Fias et al., 2013; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013), as it leads to ignoring other critical cognitive functions and processes that might play a role, such as working memory and executive functioning, including inhibition (Friso-van den Bos et al., 2013) and phonological skills (e.g., Hecht et al., 2001). Moreover, it is not unlikely that numerical magnitude processing performance itself is also determined by domain-general processes such as inhibition (Fuhs & McNeil, 2013; Gilmore et al., 2013).

Inhibition

There is substantial evidence that executive functioning plays a major role in learning during childhood (Cragg & Gilmore, 2014; St Clair-Thompson & Gathercole, 2006), also in mathematics. One executive function that has recently gained attention in the field of mathematics learning is inhibition (e.g., Gilmore et al., 2013; Szucs et al., 2013). Inhibition is an executive function that refers to *one's ability to control one's attention, behaviour, thoughts to override a strong internal predisposition or external lure and instead do what's more appropriate or needed* (Diamond, 2013, p. 137). It is an important factor, since inhibition is a domain-general skill associated with all learning-related activities (Allan et al., 2014) and inhibitory control early in life appears to be predictive of several outcomes throughout life (Moffitt et al., 2011). Children with better inhibitory control grow up to have better physical and mental health and are happier as adults (Moffitt et al., 2011). Various models of inhibition have been proposed (Diamond, 2013; Nigg, 2000). One often-used way to distinguish between different types of inhibition is by using the distinction between behavioural inhibition or response inhibition (often measured with go/no-go tasks or stop-signal tasks) and cognitive inhibition or interference control (often measured by means of stroop tasks) (Diamond, 2013).

There are several reasons to assume that inhibition is an important factor in mathematical skills. On empirical ground Koontz and Berch (1996) were the first

to suggest the idea that children with mathematical difficulties might have an inhibition impairment. Since then, various studies suggested that inhibitory control abilities are associated with performance in mathematics (e.g., Blair & Razza, 2007; Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Bull & Scerif, 2001; Espy et al., 2004; Gilmore et al., 2013; Kroesbergen, Van Luit, Van Lieshout, Van Loosbroek, & Van der Rijt, 2009; Lee et al., 2010; St Clair-Thompson & Gathercole, 2006; Thorell, 2007; see Allan et al., 2014, for a meta-analysis). Also, various studies point to poor inhibition skills as a key process in developmental dyscalculia (e.g., Bull & Scerif, 2001; De Visscher and Noël, 2013; Passolunghi & Siegel, 2004; Szucs et al., 2013), and attentional deficit hyperactivity disorder (ADHD) – which is characterized *inter alia* by poor response inhibition – has also been reported to be associated with arithmetic deficits (Kaufmann & Nuerk, 2006). On the other hand, on the neural level it has also been suggested that inhibition is associated with mathematical skills given the presumed role of the prefrontal cortex in inhibition and in mathematics (e.g., Allan et al., 2014; Willoughby, Kupersmidt, & Voegler-Lee, 2012).

The association between inhibition and mathematical skills is likely to vary depending on the mathematical skill under investigation (Cragg & Gilmore, 2014). It might be particularly prominent in arithmetic fact retrieval (Verguts & Fias, 2005), since incorrect but competing answers have to be inhibited as arithmetic facts are stored in an associative network (e.g., Ashcraft, 1987; Campbell, 1995; Jackson & Coney, 2007; McCloskey & Lindemann, 1992; Stazyk, Ashcraft, & Hamann, 1982; Verguts & Fias, 2005; Winkelman & Schmidt, 1974). Namely, when learning arithmetic facts there is considerable overlap between previously encoded items and new ones. The feature overlap between items to be remembered determines the quality of their memory trace (De Visscher & Noël, 2014b). Because of the number of features they share, arithmetic facts are particularly prone to interference (De Visscher & Noël, 2013). Thus, if a particular problem is presented, a number of neighbouring nodes in a semantic associative network are activated and have to be inhibited. For example, when retrieving the answer to 6×3 , the incorrect but competing answers to 6×2 and 6×4 , and 5×3 and 4×3 have to be inhibited. This associative confusion effect is commonly assumed to reflect interference effects (Censabella & Noël, 2007). Therefore, a lack of inhibition skills can lead to making specific errors when solving these

problems by arithmetic fact retrieval (e.g., $6 \times 3 = 24$). With good inhibitory processes children are able to inhibit irrelevant associations more quickly and thus be less likely to develop incorrect associations as they acquire additional arithmetic operations (LeFevre et al., 2013). Retrieval difficulties of children with mathematical disorders could therefore be related to inefficient inhibition of irrelevant associations (Geary, Hamson, & Hoard, 2000). Connectionist models highlight this interference in arithmetic fact retrieval caused by a densely interconnected memory structure of associations among numerical problems, operands and answers (Campbell, 1995; Verguts & Fias, 2005). The similarity of associations provokes interference and disrupts recollection (De Visscher & Noël, 2013).

Thus, based on theoretical, behavioural and neural arguments, the association between inhibition and mathematical skills and in specific arithmetic fact retrieval seemed appealing. To the best of our knowledge there are no studies examining this specific relationship.

The current study

The current study aimed to extend the abovementioned studies by focusing on one particular mathematical skill, i.e. arithmetic fact retrieval, and by investigating the unique role of inhibition (above numerical magnitude processing) in determining individual differences in arithmetic fact retrieval. By focussing on one particular mathematical skill (i.e., arithmetic fact retrieval) we aimed to extend existing research by pinpointing more carefully the association between inhibition and mathematical achievement.

A second limitation of the existing studies investigating associations between inhibition and mathematics achievement tackled by this study, is that none of these studies has included both measures of inhibition and numerical magnitude processing to explain variability in mathematical performance. As such, it remains to be determined to what extent inhibition explains additional variance in individual differences in mathematics achievement beyond what is explained by numerical magnitude processing.

To measure numerical magnitude processing, we used both symbolic and non-symbolic numerical magnitude comparison tasks, thereby enabling us to compare performance on numerical tasks with and without symbolic processing requirement (De Smedt & Gilmore, 2011). To measure inhibition skills, we used both cognitive tasks and a questionnaire, to compensate for shortcomings of both methods. More specifically, questionnaires are hampered by rater subjectivity, as opposed to cognitive tasks which are assumed to measure children's inhibition objectively and assess cognitive process involved in inhibition (Allan et al., 2014). But because cognitive tasks are administered at a single time, patterns cannot be captured and factors related but not central to inhibition (e.g., processing speed) and unrelated factors (e.g., time of testing, child fatigue) may influence results (Allan et al., 2014). Important information regarding the relation between inhibition and mathematical skills can be obtained using both cognitive or behavioural tasks and teacher reports (Allan et al., 2014), since both type of measures capture unique and important aspects of inhibition (e.g., Allan, Loningan, & Wilson, 2013; Valiente, Lemery-Chalfant, & Swanson, 2010). We used teacher ratings because teacher ratings of inhibition are more associated with measures of academic skills than parent ratings, given that teachers observe children's behaviour in relation to academic tasks (e.g., Blair & Razza, 2007). We also administered tasks to measure potentially confounding variables (i.e., intellectual ability and reading), that could explain an association between inhibition and mathematics. We assessed children in the third grade, because we wanted to study the association between arithmetic fact retrieval and inhibition with children who had already acquired a considerable number of arithmetic facts.

To examine to what extent there is an association between arithmetic fact retrieval and inhibition skills, we ran a correlation and a regression analysis. Subsequently we performed a cluster analysis, which allowed us to form subgroups based on empirically derived differences in parameters of arithmetic fact retrieval, and compare these profiles on various cognitive skills that have been associated with individual differences in arithmetic fact retrieval (e.g., inhibition, numerical magnitude processing). We examined the differences between these clusters by using an analysis of (co)variance.

Drawing on previous work, we expected that measures of inhibition skills and arithmetic fact retrieval skills would be positively correlated (the better the

inhibition skills, the better the arithmetic fact retrieval skills), even after controlling for intellectual ability. Furthermore, we expected that there would be group differences in inhibition skills between the profiles of arithmetic fact retrieval obtained by the cluster analysis. We expected that these group differences would remain significant, even after controlling for intellectual ability, numerical magnitude processing and reading, thus showing a unique role of inhibition in arithmetic fact retrieval.

Method

Participants

Initially, 102 children were invited to participate, yet the parents of seven children did not give consent. The final sample comprised 95 typically developing third-graders (51 boys, 44 girls) between 8 years 2 months and 9 years 2 months ($M = 8$ years 8 months; $SD = 4$ months). For all participants, written informed parental consent was obtained. The children were recruited from four elementary schools located in provincial towns in the middle of Flanders, Belgium and had dominantly middle- to high socio-economic background. None of them had a developmental disorder or mental retardation, nor repeated a grade.

Materials

Materials consisted of standardized tests, paper-and-pencil tasks and computer tasks designed with E-Prime 2.0 (Schneider, Eschmann, & Zuccolotto, 2002). All computer tasks were conducted on a 17-inch notebook computer. Stimuli occurred in white on a black background (Arial font, 72-point size). Response keys were always "d" (left response; labelled with a green sticker) and "k" (right response; labelled with a red sticker). The children were instructed to keep their index fingers on both keys during the task and to perform both accurately and fast. Both accuracy and reaction time (ms) were registered by the computer.

Numerical magnitude processing.

To assess children's numerical magnitude processing we used a symbolic and a non-symbolic numerical magnitude comparison task consisting of Arabic digits and dot arrays, respectively. These tasks were the same as in De Smedt and Gilmore (2011).

The tasks consisted of comparing two simultaneously presented numerical magnitudes arranged on either side of the centre of the screen. The children had to select the numerically larger magnitude by pressing the key on the side of the larger numerical magnitude. The stimuli in both tasks comprised all combinations of numerosities 1 to 9, yielding 72 trials for each task. Three practice trials were presented for each task. Per task, the stimuli were randomly divided into two blocks and children were given short breaks between blocks. Each trial started with a 200ms fixation point in the centre of the screen and after 1000ms the stimulus appeared. In the symbolic task, stimuli remained visible until response. In the non-symbolic task, stimuli disappeared after 840ms in order to avoid counting of the dots. The position of the largest numerosity was counterbalanced. The non-symbolic stimuli were generated with the MATLAB script provided by Piazza, Izard, Pinel, Le Bihan and Dehaene (2004) and were controlled for non-numerical parameters (i.e., density, dot size and total occupied area). On half of the trials dot size, array size and density were positively correlated with number, and on the other half they were negatively correlated. These visual parameters were manipulated to ensure that children could not reliably use these non-numerical cues or perceptual features to make a correct decision.

Inhibition.

To assess children's inhibition skills, we conducted both cognitive tasks (i.e., stroop tasks) and a questionnaire for the teacher.

Stroop tasks. To assess children's inhibition skills at the cognitive level, we used two measures of interference control, i.e. Stroop tasks (MacLeod, 1991), in which the processing of possibly interfering information has to be inhibited. We administered a numerical (counting) and a non-numerical (colour-word) variant of the Stroop task (van der Sluis, de Jong, & van der Leij, 2004). Both tasks involved a baseline and an interference condition and both conditions were preceded by 16

practice trials to ensure that the children understood the task. In the baseline condition the children had to name the number of objects in the numerical version (e.g., how many in $\Delta\Delta\Delta$) and name colours (i.e., coloured rectangles) in the non-numerical version. In the interference condition, the children had to name the number of objects (e.g., how many digits in 222) in the numerical version and name the ink of a colour word (e.g., blue written in red ink) in the non-numerical version. Each task included four stimuli that were repeated 10 times. All stimuli were presented on a paper, with five lines of eight stimuli. Task administration was the same in both conditions. The child had to name all the stimuli, while the experimenter registered accuracy and time to name the entire sheet.

BRIEF. The inhibition subscale of the Behavior Rating Inventory of Executive Function or BRIEF (Smidts & Huizinga, 2009) version for teachers was used to collect behavioural data of children's inhibition skills. The BRIEF is a standardized questionnaire that consists of 75 items that describe executive functioning behaviour, divided in eight subscales (e.g., inhibition, cognitive flexibility, working memory). We used the 10 items of the inhibition subscale (Cronbach $\alpha = .94$) as a behavioural measure of inhibition (e.g., 'Has difficulties controlling his/her behaviour'). The teacher rated every item for every child on a 3-point scale (never – sometimes – always). The answer 'never' scored 1 point, 'sometimes' 2 points and 'always' 3 points. The score consisted of the sum of the points on the 10 items (max = 30). Higher scores indicated more teacher-reported difficulties in inhibition.

Arithmetic fact retrieval.

Arithmetic fact retrieval was assessed by means of two verification tasks and the Tempo Test Arithmetic (de Vos, 1992).

Verification tasks. The children conducted two single-digit arithmetic verification tasks on the computer: one addition task and one multiplication task. Stimuli were selected from a standard set of single-digit arithmetic problems (LeFevre, Sadesky, & Bisanz, 1996), which excludes tie problems (e.g., $4 + 4$) and problems containing 0 and 1 as an operand or answer. The addition items comprised all combinations of the numbers 2 to 9 ($n = 28$) and each item was once presented with the correct answer, once with an incorrect answer, yielding

56 trials. The multiplication items consisted of all items with a product smaller or equal to 25 ($n = 30$), because these small problems are more likely to be solved by direct retrieval from long-term memory (e.g., Campbell & Xue, 2001). Each item was once presented with the correct answer, once with an incorrect answer, yielding 60 trials. The position of the numerically largest operand was counterbalanced. An equal number of false and correct items was presented. The false solutions in the addition task were created by adding or subtracting 1 or 2 to the solution. The false solutions in the multiplication task were table related ($n = 10$; e.g., $6 \times 3 = 24$), the answer of the corresponding addition ($n = 10$; e.g., $8 \times 2 = 10$) or unrelated ($n = 10$; e.g., $8 \times 3 = 25$). Half of the false solutions were numerically larger than the correct answer. Each task was preceded by eight practice trials to familiarize the child with the task requirements. A list of the items can be found in the Appendix. Each trial started with a 250ms fixation point in the centre of the screen and after 750ms the stimulus appeared. The stimuli remained visible until response. The children had to indicate if the presented solution for the problem was correct (by pressing the left response key, labelled with a green sticker) or false (by pressing the right response key, labelled with a red sticker). For each task, stimuli were presented randomly divided into two blocks and children were given short breaks between blocks.

Tempo Test Arithmetic. The Tempo Test Arithmetic (Tempo Test Rekenen, TTR; de Vos, 1992) is a standardized, paper-and-pencil test that measures speeded arithmetic. Four mathematical operations are tested: addition, subtraction, multiplication, and division. For each operation 40 problems of increasing difficulty were presented in one column. In one additional column 40 problems with mixed operations were presented. This resulted in a set of 200 basic arithmetic problems presented in five columns. For each column the children were instructed to solve as many problems as possible within a one-minute period. The number of correctly solved problems within the time-limit constituted their score. Their total score (i.e., the sum of all columns) was calculated.

General mathematics achievement.

Children's general mathematics achievement was assessed using a curriculum-based standardized achievement test for mathematics from the

Flemish Student Monitoring System (Leerlingvolgsysteem, LVS; Dudal, 2000). We used the test for the Middle Third Grade (Cronbach $\alpha = .90$). This untimed test consists of 60 items covering various aspects of mathematics such as number knowledge, understanding of operations, simple arithmetic, multi-digit calculation, word problem solving, measurement and geometry. The score on this test was the number of correctly solved problems.

Reading ability.

Reading ability was assessed by the standardised Dutch One-Minute Test (Eén-Minuuut Test, EMT), Version A (Brus & Voeten, 1995), which measures word decoding, and the Klepel, version A (van den Bos, Spelberg, Scheepstra, & de Vries, 1994), which measured pseudoword decoding. For the EMT, children were given a list of 116 unrelated words of increasing difficulty. Within a time-period of one minute they had to read them as accurately and quickly as possible. For the Klepel a list of 116 nonwords of increasing difficulty was presented. The children were given two minutes to read them as accurately and quickly as possible. Both tests combined speed and accuracy into one index score. The total score was the number of words read correctly within the time-limit across the two tests.

Control measures.

Intellectual ability. Raven's Standard Progressive Matrices (Raven, Court, & Raven 1992) was used as a measure of intellectual ability (Cronbach $\alpha = .88$). The children were administered 60 multiple-choice items where they had to complete a pattern. The raw score was the number of correct answers within 40 minutes. For each child a standardized score ($M = 100$, $SD = 15$) was calculated.

Motor reaction time. A motor reaction time task was included as a control for children's response speed on the keyboard. This task was the same as in De Smedt and Boets (2010). Two shapes, one of which was filled, were simultaneously displayed, one on the left and one on the right of the computer screen. The children had to press the key corresponding to the side of the filled figure. All shapes were similar in size. The administered shapes were circle, triangle, square, star and heart. Each shape occurred four times filled and four times non-filled. This resulted

in 20 trials. Three practice trials were included to familiarize the child with the task. The position of the correct answer (filled shape) was counterbalanced. Each trial started with a 200ms fixation point in the centre of the screen and after 1000ms the stimulus appeared. The stimuli remained visible until response.

Procedure

All children were tested at their own school during regular school hours. They all completed three sessions: an individual session (20 minutes), a session in small groups of four children (45 minutes) and a group-administered session (60 minutes). The individual session and small group session with the computer took place in a quiet room. All children were tested in the middle of the third grade. They all completed the tasks in the same order. The children first completed the numerical and non-numerical stroop tasks and then the EMT and Klepel. After that they completed the computer tasks: first the motor reaction task, then the addition verification task and multiplication verification task, and finally the symbolic and non-symbolic comparison tasks. In the last session they completed the Tempo Test Arithmetic and the Raven's Standard Progressive Matrices.

Results

In the stroop and computerized tasks, accuracy was very high (see Appendix). We therefore combined for all these tasks the accuracy and response times into one score, by dividing response time by the accuracy. This index was included in all subsequent analyses.

We verified if the task design of the stroop task had worked by comparing the baseline and interference condition. The interference condition was executed significantly more slowly than the baseline condition for the numerical ($t(94) = -30.53, p < .01$) and non-numerical ($t(92) = -23.92, p < .01$) task, which indicates that the task manipulation worked.

Descriptive analyses

The means, standard deviations and ranges for all administered measures are displayed in Table 1. The data were well distributed and there were no floor- or ceiling effects.

Table 1

Descriptive statistics of the administered measures.

	<i>N</i>	<i>M</i>	<i>SD</i>	Range
Numerical magnitude processing				
<i>Symbolic</i>	94	904.42	227.86	[602.47 – 1583.62]
<i>Non-symbolic</i>	94	791.55	191.39	[531.82 – 1439.98]
Inhibition				
<i>Numerical Stroop</i>				
Baseline condition	95	28.71	4.52	[19.25 – 43.93]
Interference condition	95	49.65	7.74	[34.63 – 74.20]
<i>Non-numerical Stroop</i>				
Baseline condition	93	32.35	6.09	[19.04 – 55.66]
Interference condition	93	58.17	13.09	[30.22 – 104.63]
<i>BRIEF</i>	94	13.91	4.43	[10 – 27]
Arithmetical fact retrieval				
<i>Tempo Test Arithmetic</i>	94	86.90	17.08	[39 – 122]
<i>Verification tasks</i>				
Addition	95	4138.18	1378.59	[2156.59 – 8788.95]
Multiplication	92	3570.35	1787.55	[1558.79 – 9808.52]
General mathematics achievement	94	45.10	8.50	[23 – 59]
Reading	95	46.99	12.78	[21.50 – 77.00]
Control Measures				
<i>Raven</i>	93	109.35	10.62	[72 – 130]
<i>Motor reaction time</i>	95	620.32	179.45	[358.89 – 1158.05]

Note. Reaction times in the stroop tasks are measured in seconds. All other reaction times are measured in milliseconds.

Correlational analyses

Pearson correlation coefficients were calculated to examine the associations between the different variables under study. Table 2 shows partial correlation coefficients, controlled for intellectual ability. Some correlations were controlled for additional variables (see note). Correlations with computerized tasks were controlled for motor reaction time, and correlations with the stroop tasks were controlled for the baseline conditions.

Table 2

Partial correlation coefficients between the administered measures, controlled for intellectual ability.

	1	2	3	4	5	6	7	8	9	10
1. Symbolic NMP										
2. Non-symbolic NMP	.47 ^{a**}									
3. Numerical Stroop	.18 ^b	.21 ^b								
4. Non-numerical Stroop	.21 ^c	.10 ^c	.29 ^{d**}							
5. BRIEF	.20 ^a	.16 ^a	-.08 ^e	.11 ^f						
6. Tempo Test Arithmetic	-.32 ^{a**}	-.14 ^a	-.09 ^e	-.15 ^f	.05					
7. Addition verification	.59 ^{a**}	.19 ^a	.15 ^b	.10 ^c	.08 ^a	-.53 ^{a**}				
8. Multiplication verification	.57 ^{a**}	.23 ^{a*}	.05 ^b	.14 ^c	.09 ^a	-.45 ^{a**}	.78 ^{a**}			
9. General mathematics achievement	-.13 ^a	.06 ^a	-.05 ^e	-.15 ^f	.05	.37 ^{**}	-.41 ^{a**}	-.19 ^a		
10. Reading	-.22 ^{a*}	-.17 ^a	-.20 ^e	-.34 ^{a**}	-.03	.38 ^{**}	-.27 ^{a*}	-.21 ^a	.18	
11. Fact retrieval	-.32 ^{a**}	-.14 ^a	-.07 ^e	-.20 ^f	-.06	.98 ^{**}	-.53 ^{a**}	-.45 ^{a**}	.36 ^{**}	.38 ^{**}

Note. NMP = numerical magnitude processing.

^a Controlled for motor reaction time; ^b controlled for motor reaction time and baseline numerical stroop; ^c Controlled for motor reaction time and baseline non-numerical stroop; ^d Controlled for baseline numerical stroop and non-numerical stroop; ^e Controlled for baseline numerical stroop; ^f Controlled for baseline non-numerical stroop.

* $p < .05$; ** $p < .01$.

All experimental measures that were thought to measure the same underlying component – i.e. the numerical magnitude processing tasks, the inhibition tasks and the arithmetic fact retrieval tasks – were significantly correlated with each other, except for the behavioural measure of inhibition (i.e., BRIEF questionnaire), which was not significantly correlated with the two cognitive measures of inhibition

(i.e., Stroop tasks). Because of the significant correlations between the arithmetic fact measures (all $|rs| > .45$, $ps < .01$), we combined these scores (i.e., Tempo Test Arithmetic and both verification tasks) into one composite score (i.e., *fact retrieval*; consisting of the mean of all arithmetic fact retrieval tasks scores) to improve clarity. This score was included in all subsequent analyses. It is important to point out that when the pattern of correlations was investigated for each arithmetic fact measure separately, results were very similar.

Symbolic numerical magnitude processing was significantly correlated with fact retrieval, indicating that children with better symbolic numerical magnitude processing skills showed better arithmetic fact retrieval performance. There was no significant association between non-symbolic numerical magnitude processing and arithmetic fact retrieval.

There was no significant correlation of the behavioural measure of inhibition (BRIEF) with any other measure. We found no significant correlations between the stroop tasks and arithmetic fact retrieval.

Arithmetic fact retrieval was significantly correlated with general mathematics achievement and with reading.

Regression analyses

Regression analyses were calculated to assess the amount of unique variance in arithmetic fact retrieval that was explained by inhibition and numerical magnitude processing. The assumptions for the analysis were met (see Appendix).

We tested four models: two models, one with the numerical stroop task and one with the non-numerical stroop task, for each measure of numerical magnitude processing. The results of the four regression models are presented in Table 3. Model 1 and Model 2 indicated that symbolic numerical magnitude processing significantly predicted fact retrieval ($ps < .05$), even when controlling for each inhibition task. Inhibition skills did not significantly predict arithmetic fact retrieval above symbolic numerical magnitude processing ($ps > .05$). Model 3 and 4 indicated that neither non-symbolic numerical magnitude processing, nor inhibition significantly predicted arithmetic fact retrieval ($ps > .05$).

Table 3

Regression analysis of fact retrieval.

Variable	Beta	t	Unique R ²
Model 1			
Motor reaction time	.10	.89	0.009
Intellectual ability	-.01	-.08	0.001
Symbolic NMP	-.30	-2.64**	0.073
Baseline condition numerical stroop	-.19	-1.62	0.028
Interference condition numerical stroop	-.04	-.29	0.001
Model 2			
Motor reaction time	.10	.87	0.008
Intellectual ability	.03	.25	0.006
Symbolic NMP	-.30	-2.60*	0.071
Baseline condition non-numerical stroop	-.06	-.42	0.002
Interference condition non- numerical stroop	-.19	-1.39	0.02
Model 3			
Motor reaction time	.04	.35	0.001
Intellectual ability	-.02	-.14	0.001
Non-symbolic NMP	-.10	-.85	0.014
Baseline condition numerical stroop	-.23	-1.80	0.036
Interference condition numerical stroop	-.05	-.38	0.002
Model 4			
Motor reaction time	.03	.28	0.001
Intellectual ability	.03	.25	0.008
Non-symbolic NMP	-.11	-.96	0.011
Baseline condition non-numerical stroop	-.05	-.36	0.001
Interference condition non- numerical stroop	-.24	-1.72	0.033

Note. NMP = numerical magnitude processing.

* $p < .05$; ** $p < .01$

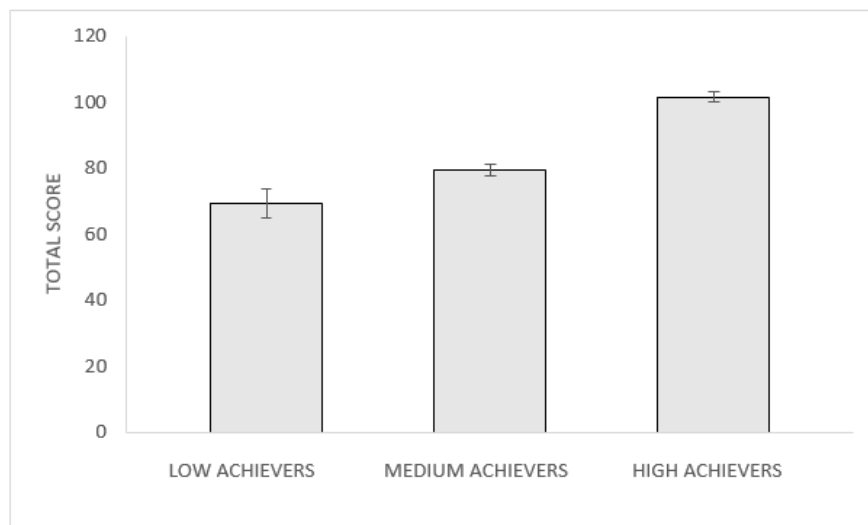
Cluster analyses

It is possible that inhibition does not show a continuous association with arithmetic fact retrieval, but that it has a different role in different subgroups of arithmetic fact retrieval skills. We therefore distinguished such subgroups in our data and examined their differences on inhibition, numerical magnitude processing, general mathematics achievement and reading. Different from a theory-driven top-down approach with a priori cut-off criteria to define subgroups, we used a data-driven bottom-up approach by using a K-means clustering approach (Wu, 2012) to delineate groups of participants in terms of their arithmetic fact retrieval skills.

We used the scores on the Tempo Test Arithmetic and both verification tasks to delineate clusters based on arithmetic fact retrieval skills. Children who lacked one of the arithmetic fact retrieval scores were not included in the cluster analysis, so in total 91 participants were included. Three groups were obtained, respectively with children that could be characterized the lowest ($n = 14$), medium ($n = 39$) and highest ($n = 38$) achievers on arithmetic fact retrieval. The arithmetic fact retrieval profiles of the obtained clusters are displayed in Figure 1, Figure 2 and Figure 3. The means and standard deviations of the clusters for the administered measures are displayed in Table 4.

Figure 1

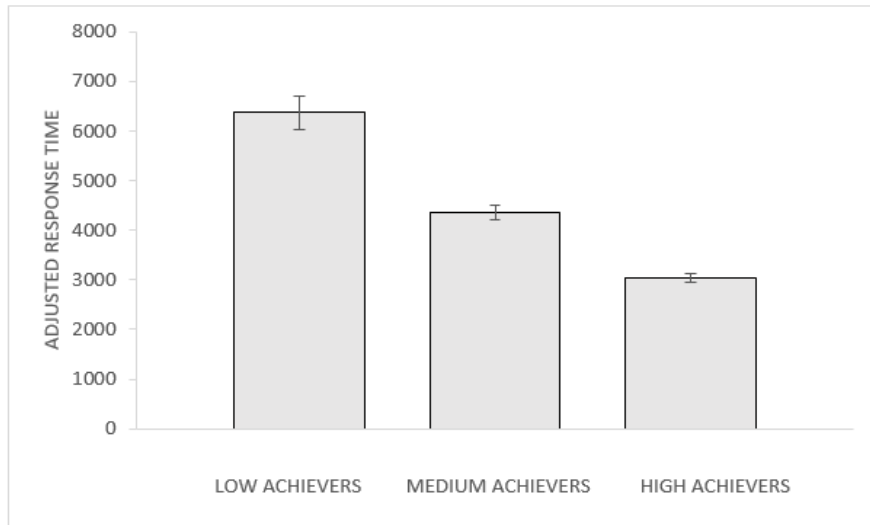
Means and error bars of the three clusters on the Tempo Test Arithmetic.



Note. Error bars represent 1 standard error of the mean.

Figure 2

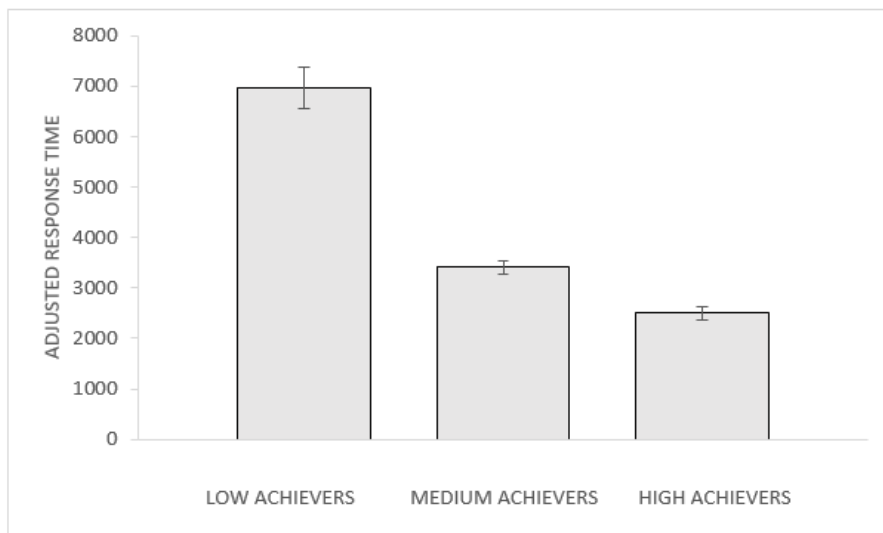
Means and error bars of the three clusters on the addition verification task.



Note. Error bars represent 1 standard error of the mean; Adjusted response time = response time / accuracy.

Figure 3

Means and error bars of the three clusters on the multiplication verification task.



Note. Error bars represent 1 standard error of the mean; Adjusted response time = response time / accuracy.

Table 4

Descriptive statistics of the identified clusters.

Variable	Low achievers		Medium achievers		High achievers	
	M	SD	M	SD	M	SD
Symbolic NMR	1183.05	228.93	895.71	197.93	806.84	173.86
Non-symbolic NMR	856.41	202.04	831.63	220.32	731.51	144.04
BC numerical stroop	32.04	4.93	28.94	4.44	27.38	3.77
IC numerical stroop	54.44	6.64	49.97	7.67	47.73	7.83
BC non-numerical stroop	33.44	5.74	32.91	5.95	31.44	6.48
IC non-numerical stroop	63.04	14.94	57.92	10.96	55.64	13.88
BRIEF	14.86	4.80	13.77	4.64	13.76	4.18
General mathematics achievement	43.43	9.21	42.59	7.73	48.27	8.43
Reading	41.14	10.79	45.42	12.91	50.95	12.29
Intellectual ability	41.08	6.78	37.67	5.93	39.05	7.94
Motor reaction time	624.32	97.48	630.84	180.81	615.30	210.12

Note. NMR = numerical magnitude representation; BC = baseline condition; IC = interference condition.

The clusters did not differ in terms of age, $F(2, 88) = 0.88, p = .42$, sex, $\chi^2(1) = 1.18, p = .56$, or intellectual ability, $F(2, 87) = 1.23, p = .30$.

To check whether the differences between the clusters on the administered variables were significant, we ran an analysis of variance (ANOVA) with the obtained clusters as a between-subjects factor and numerical magnitude processing, inhibition, general mathematics achievement and reading as dependent variables. Bonferroni adjustments were used for post hoc t -tests. Partial eta-squared (η_p^2) values were calculated as a measure of effect size.

The analysis of variance revealed a significant effect of cluster on numerical magnitude processing, general mathematics achievement and reading (Table 5). There were no other significant differences ($F_s < 1.21$).

Table 5

Differences between the arithmetic fact retrieval clusters.

Variable	<i>df</i>	<i>F</i>	<i>p</i>	η_p^2
Symbolic NMP ^a	2, 85	22.48**	0.000	0.35
Non-symbolic NMP ^a	2, 85	3.35*	0.04	0.07
Interference condition numerical stroop ^b	2, 87	0.87	0.422	0.02
Interference condition non-numerical stroop ^c	2, 85	1.21	0.303	0.03
BRIEF	2, 83	0.35	0.703	0.01
General mathematics achievement	2, 83	4.68*	0.012	0.10
Reading	2, 83	5.61**	0.005	0.12

Note. NMP = numerical magnitude processing.

^a Controlled for motor reaction time; ^b Controlled for baseline condition numerical stroop; ^c Controlled for baseline condition non-numerical stroop.

* $p < .05$; ** $p < .01$.

Post-hoc analyses demonstrated that on the symbolic numerical magnitude comparison task the lowest achievers significantly differed from the medium and high achievers ($p_s < .001$), but the high and medium achievers did not significantly

differed from each other ($p = .14$). Although the analysis of variance showed cluster differences in non-symbolic numerical magnitude comparison skills, they did not differ when compared pairwise (low versus medium, $p = 1.00$, low versus high, $p = .11$, medium versus high, $p = .10$). On general mathematics achievement high and medium achievers differed significantly ($p < .05$), but there were no significant differences between the low and high ($p = .17$) and low and medium ($p = 1.00$) achievers. On reading the low and high achievers, and medium and high achievers differed significantly ($ps = .02$), in contrast with the low and medium achievers who did not differ significantly ($p = 1.00$).

To answer our second research question, i.e. to check whether the differences between the clusters on numerical magnitude processing and inhibition remained significant after controlling for each other, we conducted an analysis of covariance (ANCOVA). This analysis showed that the difference between the clusters on symbolic numerical magnitude processing remained significant, even after controlling for the numerical stroop task ($F(2,84) = 21.56, p < .001$) and for the non-numerical stroop task ($F(2,82) = 20.83, p < .001$). The difference also remained significant after controlling for reading and general mathematics achievement ($F(2,83) = 18.29, p < .001$), on which the clusters differed significantly.

The differences between the clusters on non-symbolic numerical magnitude processing did not remain significant after controlling for the numerical stroop task ($F(2,84) = 2.66, p = .076$). However, they remained significant after controlling for the non-numerical stroop task ($F(2,82) = 3.72, p = .028$) and for reading and general mathematics achievement ($F(2,83) = 3.81, p = .026$).

The differences between the clusters on inhibition were not significant ($ps > .05$) when controlling for numerical magnitude processing.

Discussion

Understanding which cognitive processes underlie individual differences in mathematical skills is a necessary prerequisite to design validated diagnostic tests and appropriate interventions. Numerical magnitude processing has been pointed out as an important factor of these individual differences (De Smedt et al., 2013), but the predominant focus on numerical magnitude processing has recently been

criticized (Fias et al., 2013). On the other hand, several studies provided evidence for an association between inhibition and individual differences in mathematics achievement (Allan et al., 2014, for a meta-analysis). However, it is currently not clear how this association with inhibition occurs in specific aspects of mathematical achievement, such as arithmetic fact retrieval. Importantly, it needs to be verified whether the association between inhibition and mathematics achievement remains when other crucial factors, that have been shown to contribute to individual differences in mathematical competence (e.g., numerical magnitude processing), are controlled for. This study tried to address these questions by examining the association between inhibition and arithmetic fact retrieval and by investigating the unique contribution of inhibition to arithmetic fact retrieval above numerical magnitude processing. We used a correlational design and a clustering approach. Based on previous findings, we hypothesized that inhibition and arithmetic fact retrieval skills would correlate positively. We further verified if there was a unique contribution of inhibition in arithmetic fact retrieval after controlling for numerical magnitude processing.

The present study failed to observe a significant association between inhibition and arithmetic fact retrieval. Consequently, our results did not reveal a unique contribution of inhibition to arithmetic fact retrieval above numerical magnitude processing. These findings are not in line with the theoretically postulated association between arithmetic fact retrieval and inhibition (e.g., Barrouillet, Fayol, & Lathulière, 1997; Geary, 2010; Geary et al., 2000; Geary et al., 2012; Verguts & Fias, 2005) and contradict previous studies that showed an association between inhibition and mathematics achievement (e.g., Bull & Scerif, 2001; LeFevre et al., 2013; Szucs et al., 2013).

This inconsistency can be explained by various factors. Firstly, it could be due to differences in the measurement of inhibition. LeFevre and colleagues (2013) used the broad concept *executive attention*, which they defined as the common aspects of executive function and working memory that are necessary in many complex cognitive tasks, including inhibition of competing responses, goal maintenance, and response selection. These authors used span tasks and a colour trail test to measure executive attention. These measures tap into broader executive functions rather than inhibition per se, which might explain the differences between their findings and the current study. Geary and colleagues

(2012) measured inhibition by counting the number of intrusion errors on an arithmetic task. As a result, these authors derived an index of inhibition from the mathematical task under study, but they did not use an independent measure of inhibition as in the current study. This measurement difference might again explain differences in results.

Secondly, the inconsistency might be due to the differences in samples that were used. Szucs and colleagues (2013) showed that children with dyscalculia performed significantly more poorly on several inhibition measures (including the number stroop), pointing to an inhibition deficit in these children. However, the current sample comprised typically developing children. Indeed, it is possible that inhibition has a specific role in arithmetic fact retrieval in children with mathematical disabilities but not in typically developing children, which could explain the different results between Szucs and colleagues (2013) and the current study.

The present findings are in agreement with the study of van der Sluis and colleagues (2004), who found that children with mathematical disabilities did not differ from the control group on the same inhibition tasks as in our study. Our results also converge with Censabella and Noël (2007), who found no significant differences in inhibition between children with mathematical disabilities and controls. However, as noted, it is not clear whether inhibition plays a (dis)continuous role in typically and atypically developing children. The current data therefore extend those of van der Sluis and colleagues (2004) and Censabella and Noël (2007) by showing no significant association in typically developing children. Importantly, it is unclear whether the children with mathematical disabilities in van der Sluis and colleagues (2004) and Censabella and Noël (2007) had specific problems with arithmetic fact retrieval. For example, Censabella and Noël (2007) stated that, although assumed, their sample of children with mathematical disabilities did not have difficulties in arithmetic fact retrieval per se. Therefore, these studies should be replicated in samples of children with specific difficulties in arithmetic fact retrieval. In the present study, we delineated different clusters based on measures of fact retrieval, including a poor fact retrieval cluster, to verify this hypothesis. Our results did not reveal poorer inhibition in this poor fact retrieval cluster. However, the children in this cluster might not be sufficiently

low achieving to detect an association between inhibition and difficulties in arithmetic fact retrieval. Further research should examine this hypothesis.

The present study also investigated the association between numerical magnitude processing and arithmetic fact retrieval. We used both symbolic and non-symbolic tasks to verify whether the access to numerical magnitudes from symbolic digits or numerical magnitude processing per se is related to arithmetic fact retrieval (see De Smedt et al., 2013, for a discussion). In line with De Smedt et al. (2013) and Vanbinst et al. (2012), Vanbinst, Ceulemans, et al. (2015) and Vanbinst, Ghesquière, et al. (2015), the importance of symbolic numerical magnitude processing in arithmetic fact retrieval was supported by our results. More specifically, we found a unique association between symbolic numerical magnitude processing and fact retrieval. The cluster analysis revealed that the three arithmetic fact retrieval clusters differed in symbolic numerical magnitude processing and this difference remained significant, even after controlling for inhibition, reading and general mathematics achievement. Thus, children with better symbolic numerical magnitude processing skills showed better arithmetic fact retrieval performance.

The results on the association between arithmetic fact retrieval and non-symbolic numerical magnitude processing were not so univocal, which is in line with De Smedt and colleagues (2013) who concluded that the data on the association between non-symbolic numerical magnitude processing and mathematics achievement have been inconclusive so far. We observed no significant correlation between non-symbolic numerical magnitude processing and arithmetic fact retrieval, which is in line with Holloway and Ansari (2009), Sasanguie et al. (2013), Vanbinst et al. (2012). However, our cluster analysis showed differences between the arithmetic fact retrieval clusters on the non-symbolic comparison tasks, but these differences were not significant in the post hoc analysis, suggesting that the effect of non-symbolic magnitude processing on arithmetic fact retrieval is only weak.

Recently, Gilmore and colleagues (2013), and Fuhs and McNeil (2013) suggested that performance on the non-symbolic numerical magnitude processing task is determined by domain-general processes such as inhibition. We also explored this issue in our data. However, we failed to observe an association

between our measures of inhibition and non-symbolic numerical magnitude processing.

Limitations and future directions

There are a number of limitations of the present study that should be kept in mind when evaluating its findings. Firstly, our measure of arithmetic fact retrieval might have been too coarse. We used single-digit arithmetic tasks, but we did not directly examine the strategies that children applied during the administration of our tasks. Therefore, our measure of arithmetic fact retrieval might have included not only retrieval but also procedural strategies. Future studies should take into account the different strategies children apply during arithmetic problem solving. This could be done by asking them on a trial-by-trial basis to verbally report how they solved the problem (e.g., Geary et al., 2012; Siegler, 1987; Vanbinst et al., 2012; Vanbinst, Ghesquière, et al., 2015;). Another option would be to administer a forced retrieval task (e.g., De Visscher & Noël, 2014b; Geary et al., 2012), for example by reducing the presentation time and restricting the response interval, in which children are forced to use retrieval strategies. Such a forced retrieval task also increases the likelihood of errors, some of which might be an indicator of difficulties in inhibitory control (e.g., table-related errors; e.g., Barrouillet et al., 1997; Geary et al., 2012)

Secondly, the present study only included one specific measure of inhibition, i.e. Stroop tasks, yet inhibition is not a unitary construct, but a family of functions (Censabella & Noël, 2007; Dempster & Corkill, 1999; Harnishfeger, 1995; Hasher, Zacks, & May, 1999; Nigg, 2000; Shilling, Chetwynd, & Rabbitt, 2002). Different aspects of inhibitory control are dissociable from each other at both the behavioural and the neural level (Diamond, 2013). One often-used distinction between different types of inhibition is between behavioural inhibition and cognitive control (Diamond, 2013). Different tasks (e.g., Stroop tasks, Flanker task, go/no-go task, stop-signal task) are used to measure these different types of inhibition, and their association with arithmetic fact retrieval might be different. Future studies should therefore use a combination of different measurements of inhibition. It is important to note that we also included a teacher questionnaire, the BRIEF (Smidts & Huizinga, 2009) to measure inhibition skills in children, but this measure was not associated with arithmetic fact retrieval. Although the inhibition subscale of the BRIEF is known to be reliable, the use of 10 items might

have been limited to capture sufficient intersubject variability. On the other hand, the current data are in line with several studies who failed to observe a significant correlation between ratings (e.g., BRIEF questionnaire) and performance-based measures (e.g., stroop tasks), suggesting these measures assess different aspects of the same underlying construct (Mahone et al., 2002; McAuley, Chen, Goos, Schachar, & Crosbie, 2010; Toplak, West, & Stanovich, 2013).

Thirdly, there are several versions of the Stroop task (e.g., animal stroop (Szucs et al., 2013), spatial stroop (Diamond, 2013), numerical stroop (Bull & Scerif, 2001)), and, in particular, of the numerical stroop task. We used a counting stroop task as numerical stroop task. The difference between the different versions of the tasks lies in the information that has to be inhibited. In our numerical stroop task (i.e., Counting Stroop Task), the number represented by the digits had to be inhibited in favour of the quantity of digits in the array. On the other hand, in the Number Stroop (e.g., Kaufmann & Nuerk, 2006) magnitudes of two one-digit numbers who differ in physical size are compared and participants have to inhibit the irrelevant physical size in favour of the numerical magnitude of the digits (or vice versa). Because these tasks contain different kinds of numerical information that needs to be inhibited, they may be differently related to arithmetic fact retrieval.

Fourthly, Censabella and Noël (2007) stated that active inhibition processes might not be involved in arithmetic fact retrieval. This is because the activation-based interference does not necessarily require active inhibitory mechanisms, as the reduced associative strength of the correct answer is sufficient to account for weaker performances. This interference is of a passive kind due to the 'overfacilitation' of competitors. This is in line with De Visscher and Noël's (2014a; 2014b) notion of hypersensitivity-to-interference, which states that the similarity between arithmetic facts provokes interference, and learners who are hypersensitive to interference will therefore encounter difficulties in storing arithmetic facts in long-term memory. De Visscher and Noël (2014b) found clear detrimental effects of interference in multiplication facts storing. For example, the degree of interference influences the performance across multiplications and determines part of the individual differences in multiplication. De Visscher and Noël (2014a) found that children with low arithmetic fluencies experience hypersensitivity-to-interference in memory compared to children with typical

arithmetic fluencies. It turns out that this sensitivity-to-interference-parameter does not correlate with inhibition (i.e., Colour-Word Stroop task) (De Visscher and Noël, 2014b), suggesting that sensitivity to interference does not correspond to dominant response inhibition capacities. We did not include such tasks of passive inhibition in our study, an issue that should be considered in future research.

Fifthly, we only investigated children in third grade, which might explain why we did not find an association between inhibition skills and arithmetic fact retrieval. Although children in third grade have already acquired a considerable number of arithmetic facts, there is still room for improvement in automatizing these facts. Over time, children rely less on effortful and time-consuming procedural strategies (e.g., finger counting, decomposition), but they increasingly use direct and fast retrieval of arithmetic facts (e.g., Bailey, Littlefield, & Geary, 2012; Barrouillet, Mignon, & Thevenot, 2008; Geary, 2013; Siegler, 1996; Vanbinst, Ceulemans, et al., 2015). Through the course of primary school, problem-answer associations become stronger, and more efficient arithmetic fact retrieval arises (Vanbinst, Ceulemans, et al., 2015). On the other hand, inhibitory control also continues to mature through the course of primary school (Luna, 2009; Luna, Garver, Urban, Lazar, & Sweeney, 2004). Therefore, it would be interesting for future studies to investigate the association between inhibition and fact retrieval in older children.

Sixthly, the present study only comprised typically developing children. It might be that the association between inhibition and arithmetic fact retrieval only is observed in the context of atypical development of arithmetic fact retrieval and/or of atypical development of inhibition. Future studies should examine the association between inhibition and arithmetic fact retrieval skills in atypical groups, such as children with arithmetic fact retrieval deficits and children with ADHD – who are known to have deficits in (response) inhibition (e.g., Bayliss & Roodenrys, 2000; Kaufmann & Nuerk, 2006).

Seventhly, the use of k-means cluster analysis in the present study is not without limitation. The lack of significant differences between the low and medium achieving groups on the Tempo Test Arithmetic and the combined fact retrieval score suggests that our clusters were somewhat overlapping. Future studies should use a model-based clustering approach, which is preferred over heuristic clustering methods, as in the current study, because it provides a principled statistical approach (Banfield & Raftery, 1993). The use of model-based clustering

also allows one to use the Bayesian information criterion (BIC), which weighs model fit and model complexity, to decide how many clusters are needed to adequately describe the data.

The association between inhibition and mathematics observed in previous studies might also be explained by other factors that are associated with both individual differences in mathematics and inhibition. Potential examples include socio-economic status and home environment (Dilwordt-Bart, 2012; Sarsour et al., 2011). These factors should be considered in future studies.

Future studies should also examine the association between inhibition and arithmetic fact retrieval at the neural level. Even though this association might not be detectable at the behaviour level, it might be that it can be observed at the neural level. Indeed, neuroimaging data might generate findings that cannot be detected by behavioural data alone (De Smedt et al., 2010). Brain areas associated with inhibition (e.g., prefrontal cortex) are often found to be activated during mathematical tasks (e.g., Menon, 2015, for a review). Although many fMRI studies have pointed to prefrontal cortex control processes during arithmetic fact retrieval (Menon, 2015), there is no study that has directly investigated the overlap between these control networks and arithmetic fact retrieval. Cho and colleagues (2012) found that increased retrieval use was correlated with the dorsolateral and ventrolateral prefrontal cortex, areas that are also known to show increased activity during inhibition. The authors suggested that this increase in the lateral prefrontal cortex suggested the involvement of inhibitory processes, yet they did not directly test this hypothesis. Future studies should investigate this hypothesis with imaging studies, for example by investigating the neural overlap between an arithmetic task and an inhibition localizer task.

References

- Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., & Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: a meta-analysis. *Developmental Psychology, 50*(10), 2368-2379. doi: 10.1037/a0037493
- Allan, N. P., Lonigan, C. J., & Wilson, S. B. (2013). Psychometric evaluation of the Children's Behavior Questionnaire – Very Short Form in preschool children using parent and teacher report. *Early Childhood Research Quarterly, 28*, 302-313. doi: 10.1016/j.ecresq.2012.07.009
- Ancker, J. S., & Kaufman, D. (2007). Rethinking health numeracy: A multidisciplinary literature review. *Journal of the American Medical Informatics Association, 14*(6), 713-721.
- Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience, 9*(4), 278-291. doi: 10.1038/nrn2334
- Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In C. J. Brainerd, R. Kail, & J. Bisanz (Eds.), *Formal methods in developmental research* (pp. 302-338). New York: Springer-Verlag.
- Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The co-development of skill at and preference of use of retrieval-based processes for solving addition problems: individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology, 113*, 78-92. doi: 10.1016/j.jecp.2012.04.014
- Banfield, J. D., & Raftery, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics, 49*(3), 803-821.

- Barrouillet, P., & L epine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology*, *91*, 183-204. doi: 10.1016/j.jecp.2005.03.002
- Barrouillet, P., Fayol, M., & Lathuli re, E. (1997). Selecting between competitors in multiplication tasks: An explanation of the errors produced by adolescents with learning disabilities. *International Journal of Behavioural Development*, *21*(2), 253-275.
- Barrouillet, P., Mignon, M., & Thevenot, C. (2008). Strategies in subtraction problem solving in children. *Journal of Experimental Child Psychology*, *99*(4), 233-251. doi: 10.1016/j.jecp.2007.12.001
- Bayliss, D. M., & Roodenrys, S. (2000). Executive processing and ADHD: An application of the supervisory attentional system. *Developmental Neuropsychology*, *17*(2), 161-180. doi: 10.1207/S15326942DN1702_02
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647-663.
- Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contribution of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, *24*(3), 337-349. doi: 10.1016/j.ecresq.2009.06.001
- Brus, B. T., & Voeten, B. J. (1995). *E en minuut test vorm A en B. Verantwoording en handleiding* [One Minute Test version A and B. Justification and manual]. Nijmegen, The Netherlands: Berkhout.

- Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*(3), 273-293. doi: 10.1207/S15326942DN1903_3
- Byrnes, J. P., & Wasik, B. A. (2009). Factors predictive of mathematics achievement in kindergarten, first and third grades: An opportunity-propensity analysis. *Contemporary Educational Psychology*, *34*(2), 167-183. doi: 10.1016/j.cedpsych.2009.01.002
- Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition*, *1*(2), 121-164.
- Campbell, J. I. D., & Xue, Q. L. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*, 299-315. doi: 10.1037/0096-3445.130.2.299
- Censabella, S., & Noël, M. P. (2007). The inhibition capacities of children with mathematical disabilities. *Child Neuropsychology*, *14*(1), 1-20. doi: 10.1080/09297040601052318
- Cho, S., Metcalfe, A. W. S., Young, C. B., Ryali, S., Geary, D. C., & Menon, V. (2012). Hippocampal-prefrontal engagement and dynamic causal interactions in the maturation of children's fact retrieval. *Journal of Cognitive Neuroscience*, *24*(9), 1849-1866. doi: 10.1162/jocn_a_00246
- Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*(2), 63-68. doi: 10.1016/j.tine.2013.12.001

- De Smedt, B., & Boets, B. (2010). Phonological processing and arithmetic fact retrieval: Evidence from developmental dyslexia. *Neuropsychologia*, *48*(14), 3973-3981. doi: 10.1016/j.neuropsychologia.2010.10.018
- De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology*, *108*(2), 278-292. doi: 10.1016/j.jecp.2010.09.003
- De Smedt, B., Ansari, D., Grabner, R. H., Hannula, M. M., Schneider, M., & Verschaffel, L. (2010). Cognitive neuroscience meets mathematics education. *Educational Research Review*, *5*, 97-105. doi: 10.1016/j.edurev.2009.11.001
- De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). The relationship between symbolic and non-symbolic numerical magnitude processing skills and the typical and atypical development of mathematics: A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48-55. doi: 10.1016/j.tine.2013.06.001
- De Smedt, B., Reynvoet, B., Swillen, A., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Basic number processing and difficulties in single-digit arithmetic: Evidence from velo-cardio-facial syndrome. *Cortex*, *45*(2), 177-188. doi: 10.1016/j.cortex.2007.06.003
- De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*(4), 469-479. doi: 10.1016/j.jecp.2009.01.010
- De Visscher, A., & Noël, M. P. (2013). A case study of arithmetic facts dyscalculia caused by a hypersensitivity-to-interference in memory. *Cortex*, *49*(1), 50-70. doi: 10.1016/j.cortex.2012.01.003

- De Visscher, A., & Noël, M.P. (2014a). Arithmetic facts storage deficit: the hypersensitivity-to-interference in memory hypothesis. *Developmental Science*, *17*(3), 434-442. doi: 10.1111/desc.12135
- De Visscher, A., & Noël, M.P. (2014b). The detrimental effect of interference in multiplication facts storing: typical development and individual differences. *Journal of Experimental Psychology: General*, *143*(6), 2380-2400. doi: 10.1037/xge0000029
- de Vos, T. (1992). *Tempo-Test-Rekenen. Handleiding* [Tempo Test Arithmetic. Manual]. Nijmegen, The Netherlands: Berkhout.
- Dempster, F. N., & Corkill, A. J. (1999). Individual differences in susceptibility to interference and general cognitive ability. *Acta Psychologica*, *101*, 395-416. doi: 10.1016/S0001-6918(99)00013-X
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*, 135-168. doi: 10.1146/annurev-psych-113011-143750
- Dilworth-Bart, J.E. (2012). Does executive function mediate SES and home quality associations with academic readiness? *Early Childhood Research Quarterly*, *27*(3), 416-425. doi: 10.1016/j.ecresq.2012.02.002
- Donlan, C., Cowan, R., Newton, E., & Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition*, *103*(1), 23-33. doi: 10.1016/j.cognition.2006.02.007
- Dowker, A. (2005). *Individual differences in arithmetic. Implications for psychology, neuroscience and education*. Hove: Psychology Press.
- Dudal, P. (2000). *Leerlingvolgsysteem: Wiskunde-Toetsen 1-2-3 Basisboek* [Student monitoring system: Mathematics-Tests 1-2-3 manual]. Leuven, Belgium: Garant.

- Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology*, 91(2), 113-136. doi: 10.1016/j.jecp.2005.01.003
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, 26(1), 465-486. doi: 10.1207/s15326942dn2601_6
- Fias, W., Menon, V., & Szucs, D. (2013). Multiple components of development dyscalculia. *Trends in Neuroscience and Education*, 2(2), 43-47. doi: 10.1016/j.tine.2013.06.006
- Finnie, R., & Meng, R. (2001). Cognitive skills and the youth labour market. *Applied Economics Letters*, 8, 675-679.
- Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, 10, 29-44. doi: 10.1016/j.edurev.2013.05.003
- Fuhs, M., & McNeil, N. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. *Developmental Science*, 16(1), 136-148. doi: 10.1111/desc.12013.
- Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin*, 114(2), 345-362. doi: 10.1037/0033-2909.114.2.345
- Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, 37, 4-15. doi: 10.1177/00222194040370010201

- Geary, D. C. (2010). Mathematical disabilities: Reflections on cognitive, neuropsychological, and genetic components. *Learning and Individual Differences, 20*, 130-133. doi: 10.1016/j.lindif.2009.10.008
- Geary, D. C. (2013). Early foundations for mathematics learning and their relation to learning disabilities. *Current Directions in Psychological Science, 22*(1), 23-27. doi: 10.1177/0963721412469398
- Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77*(3), 236-263. doi: 10.1006/jecp.2000.2561
- Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of Learning Disabilities, 45*(4), 291-307. doi: 10.1177/0022219410392046
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One, 8*(6), e67374. doi: 10.1371/journal.pone.0067374
- Halberda, J., Mazocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665-668. doi: 10.1038/nature07246
- Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C.J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175-204). San Diego, CA: Academic Press.

- Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance xvii. Cognitive regulation of performance: Interaction of theory and application* (pp. 653-675). Cambridge, MA: The MIT Press.
- Hecht, S. A., Torgesen, J. K., Wagner, R. K. & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology, 79*(2), 192-227. doi: 10.1006/jecp.2000.2586
- Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*(1), 17-29. doi: 10.1016/j.jecp.2008.04.001
- Jackson, N., & Coney, J. (2007). Simple arithmetic processing: Individual differences in automaticity. *European Journal of Cognitive Psychology, 19*, 141-160. doi: 10.1080/09541440600612712
- Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*(2), 82-88. doi: 10.1016/j.lindif.2009.07.004
- Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology, 85*(2), 103-119. doi: 10.1016/S0022-0965(03)00032-8
- Kaufmann, L., & Nuerk, H. C. (2006). Interference effects in a numerical stroop paradigm in 9- to 12-year-old children with ADHD-C. *Child Neuropsychology, 12*(3), 223-243. doi: 10.1080/09297040500477483

- Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academies Press.
- Kononen, T., Salmi, P., Eklund, K., & Aro, T. (2013). Counting and RAN: Predictors of arithmetic calculation and reading fluency. *Journal of Educational Psychology, 105*, 162-175. doi: 10.1037/a0029285
- Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. *Mathematical Cognition, 2*, 1-23. doi: 10.1080/135467996387525
- Kroesbergen, E., Van Luit, J., Van Lieshout, E., Van Loosbroek, E., & Van der Rijt, B. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment, 27*(3), 226-236. doi: 10.1177/0734282908330586
- Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8-9-year-old students. *Cognition, 93*(2), 99-125. doi: 10.1016/j.cognition.2003.11.004
- Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., & Bull, R. (2010). The cognitive underpinnings of emerging mathematical skills: executive functioning, patterns, numeracy, and arithmetic. *British Journal of Educational Psychology, 82*(1), 82-99. doi: 10.1111/j.2044-8279.2010.02016.x
- LeFevre, J. A., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S. L., & Smith-Chant, B. L. (2013). The role of executive attention in the acquisition of mathematical skills for children in Grades 2 through 4. *Journal of Experimental Child Psychology, 114*(2), 243-261. doi: 10.1016/j.jecp.2012.10.005

- LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 216-230.
- Luna, B. (2009). Developmental changes in cognitive control through adolescence. *Advances in Child Development and Behavior*, *37*, 233-278.
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*(5), 1357-1372. doi: 10.1111/j.1467-8624.2004.00745.x
- Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, *30*(5), 520-540.
- MacLeod, C. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*(2), 163-203. doi: 10.1037/0033-2909.109.2.163
- Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., Singer, H. S., & Denckla, M. B. (2002). Validity of the behaviour rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*(7), 643-662. doi: 10.1016/S0887-6177(01)00168-8
- Mayes, S. D., Calhoun, S. L., Bixler, E. O., & Zimmerman, D. N. (2009). IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences*, *19*(2), 238-241. doi: 10.1016/j.lindif.2008.09.001

- McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*(3), 495-505. doi: 10.1017/S1355617710000093
- McCloskey, M., & Lindemann, A. M. (1992). Mathnet: Preliminary results from a distributed model of arithmetic fact retrieval. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 365-409). Amsterdam: Elsevier.
- Menon, V. (2015). Arithmetic in the Child and Adult Brain. In R. Cohen-Kadosh, & A. Dowker (Eds.) *The Oxford Handbook of Mathematical Cognition* (pp. 502-530). Oxford: Oxford University Press.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 2693-2698. doi: 10.1073/pnas.1010076108
- Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*(4), 490-502. doi: 10.1016/j.jecp.2009.02.003
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220-246. doi: 10.1037/0033-2909.126.2.220

- Opdenakker, M. C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, *33*(2), 179-206. doi: 10.1080/01411920701208233
- Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* London: National Research and Development Centre for Adult Literacy.
- Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology*, *88*(4), 348-367. doi: 10.1016/j.jecp.2004.04.002
- Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547-555. doi: 10.16/j.neuron.2004.10.014
- Raghubar, K., Barnes, F., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110-122. doi: 10.1016/j.lindif.2009.10.005
- Raven, J. C., Court, J. H. & Raven, J. (1992). *Standard progressive matrices*. Oxford, United Kingdom: Oxford Psychologists Press.
- Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological science*, *24*, 1301-1308. doi : 10.1177/0956797612466268
- Robinson, C. S., Menchetti, B. M. & Torgesen, J. K. (2002). Toward a two-factor theory of one type of mathematics disabilities. *Learning Disabilities Research & Practice*, *17*(2), 81-89. doi: 10.1111/1540-5826.00035

- Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., & Boyce, W.T. (2011). Family socioeconomic status and child executive functions: The roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society*, 17(1), 120-132. doi: 10.1017/S1355617710001335
- Sasanguie, D., Gobel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, 114(3), 418-431. doi: 10.1016/j.jecp.2012.10.012
- Schneider W., Eschmann, A., & Zuccolotto, A. (2002). *E-prime reference guide*. Pittsburgh, PA: Psychology Software Tools.
- Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia*, 40(6), 605-619. doi: 10.1016/S0028-3932(01)00157-9
- Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, 116(3), 250-264. doi: 10.1037/0096-3445.116.3.250
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Smidts, D. P. & Huizinga, M. (2009). *BRIEF Executieve Functies Gedragsvragenlijst: Handleiding* [BRIEF Executive Functions Behavior Inventory]. Amsterdam: Hogrefe Uitgevers.

- St Clair-Thomson, H., & Gathercole, S. (2006). Executive functions and achievement in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*(4), 745-759. doi: 10.1080/17470210500162854
- Stazyk, E. H., Ashcraft, M. H., & Hamann, M. S. (1982). A network approach to mental multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *8*(4), 320-335. doi: 10.1037/0278-7393.8.4.320
- Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80-90. doi: 10.1016/j.lindif.2008.05.004
- Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*(10), 2674-2688. doi: 10.1016/j.cortex.2013/06.007
- Thorell, L. B. (2007). Do delay aversion and executive function deficits make distinct contributions to the functional impact of ADHD symptoms? A study of early academic skill deficits. *Journal of Child Psychology and Psychiatry*, *48*(11), 1061-1070. doi: 10.1111/j.1469-7610.2007.01777.x
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*(2), 131-143. doi: 10.1111/jcpp.12001
- Valiente, C., Lemery-Chalfant, K., & Swanson, J. (2010). Prediction of kindergartners' academic achievement from their effortful control and emotionality: Evidence for direct and moderate relations. *Journal of Educational Psychology*, *102*(3), 550-560. doi: 10.1037/a0018992

- van den Bos, K. P., Spelberg, H. C. L., Scheepstra, A. J. M., & De Vries, J. R. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden* [Standardized test of reading ability Form A and B]. Nijmegen: Berkhout.
- van der Sluis, S., de Jong, P.F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology*, 87(3), 239-266. doi: 10.1016/j.jecp.2003.12.002
- Vanbinst, K., Ceulemans, E., Ghesquière, P., & De Smedt, B. (2015). Profiles of children's arithmetic fact development: A model-based clustering approach. *Journal of Experimental Child Psychology*, 133, 29-46. doi: 10.1016/j.jecp.2015.01.003
- Vanbinst, K., Ghesquière, P., & De Smedt, B. (2012). Numerical magnitude representations and individual differences in children's arithmetic strategy use. *Mind, Brain and Education*, 6(3), 129-136. doi: 10.1111/J.1751-228X.2012.01148.x
- Vanbinst, K., Ghesquière, P., De Smedt, B. (2015). Does numerical processing uniquely predict first graders' future development of single-digit arithmetic? *Learning & Individual Differences*, 37, 153-160. doi: 10.1016/j.lindif.2014.12.004
- Verguts, T., & Fias, W. (2005). Interacting neighbors: a connectionist model of retrieval in single-digit multiplication. *Memory and Cognition*, 33(1), 1-16.
- Willoughby, M. T., Kupersmidt, J. B., & Voegler-Lee, M. E. (2012). Is preschool executive function causally related to academic achievement? *Child Neuropsychology*, 18(1), 79-91. doi: 10.1080/09297049.2011.578572

- Winkelman, J. H., & Schmidt, J. (1974). Associative confusions in mental arithmetic. *Journal of Experimental Psychology*, 102, 734-736.
- Wu, J. (2012). *Advances in K-means clustering. A data mining thinking*. Berlin Heidelberg: Springer-Verlag.

Appendix

1. List of problems administered in the arithmetic verification tasks
2. Missing data
3. Supplementary results

1. List of problems administered in the arithmetic verification tasks

Addition	Multiplication
<i>Practice list</i>	<i>Practice list</i>
$2 + 2 = 4$	$4 \times 4 = 16$
$5 + 1 = 7$	$6 \times 3 = 17$
$7 + 3 = 10$	$4 \times 3 = 12$
$5 + 9 = 14$	$8 \times 2 = 16$
$8 + 7 = 17$	$2 \times 7 = 13$
$2 + 6 = 10$	$6 \times 2 = 18$
$6 + 5 = 9$	$4 \times 5 = 20$
$8 + 4 = 12$	$7 \times 3 = 24$
<i>Test list 1</i>	<i>Test list 1</i>
$3 + 2 = 6$	$6 \times 2 = 14$
$2 + 6 = 10$	$3 \times 5 = 15$
$8 + 4 = 12$	$7 \times 2 = 14$
$3 + 9 = 11$	$6 \times 3 = 18$
$5 + 4 = 7$	$2 \times 4 = 8$
$8 + 2 = 10$	$7 \times 3 = 24$
$4 + 5 = 9$	$6 \times 2 = 13$
$4 + 7 = 11$	$4 \times 3 = 12$
$6 + 9 = 15$	$6 \times 4 = 24$
$9 + 2 = 12$	$2 \times 3 = 9$
$2 + 8 = 9$	$5 \times 4 = 20$
$8 + 3 = 13$	$6 \times 2 = 12$
$5 + 3 = 8$	$3 \times 8 = 32$
$9 + 5 = 14$	$5 \times 3 = 20$
$6 + 4 = 10$	$4 \times 6 = 10$
$2 + 5 = 7$	$3 \times 7 = 21$
$9 + 8 = 15$	$2 \times 5 = 10$
$7 + 3 = 10$	$8 \times 3 = 24$
$9 + 4 = 15$	$3 \times 2 = 5$
$5 + 7 = 13$	$5 \times 2 = 15$
$6 + 2 = 8$	$2 \times 8 = 17$
$4 + 3 = 6$	$6 \times 2 = 13$
$8 + 7 = 13$	$2 \times 5 = 10$
$3 + 5 = 9$	$2 \times 3 = 6$
$7 + 6 = 11$	$2 \times 9 = 18$
$9 + 3 = 12$	$5 \times 4 = 15$
$4 + 8 = 14$	$3 \times 8 = 24$
$6 + 3 = 8$	$4 \times 3 = 7$

<i>Test list 2</i>	<i>Test list 2</i>
$9 + 7 = 16$	$6 \times 2 = 12$
$2 + 3 = 5$	$5 \times 3 = 15$
$6 + 5 = 9$	$8 \times 2 = 10$
$4 + 2 = 6$	$6 \times 3 = 24$
$8 + 5 = 12$	$4 \times 2 = 8$
$6 + 7 = 13$	$8 \times 3 = 25$
$7 + 2 = 11$	$4 \times 6 = 24$
$6 + 8 = 15$	$2 \times 7 = 9$
$5 + 9 = 16$	$3 \times 4 = 12$
$2 + 4 = 7$	$9 \times 2 = 18$
$3 + 8 = 11$	$6 \times 4 = 10$
$4 + 7 = 10$	$3 \times 7 = 20$
$8 + 6 = 14$	$4 \times 5 = 25$
$7 + 9 = 14$	$8 \times 2 = 16$
$5 + 6 = 11$	$3 \times 6 = 18$
$7 + 8 = 15$	$4 \times 3 = 12$
$4 + 6 = 11$	$7 \times 2 = 13$
$2 + 9 = 11$	$2 \times 9 = 11$
$3 + 7 = 12$	$6 \times 2 = 12$
$8 + 9 = 17$	$4 \times 3 = 13$
$3 + 6 = 9$	$2 \times 8 = 16$
$7 + 5 = 12$	$3 \times 4 = 13$
$4 + 9 = 13$	$5 \times 2 = 10$
$2 + 7 = 9$	$3 \times 6 = 9$
$5 + 2 = 6$	$4 \times 5 = 20$
$3 + 4 = 7$	$7 \times 3 = 21$
$9 + 7 = 16$	$3 \times 2 = 6$
$5 + 9 = 16$	$2 \times 4 = 6$
	$3 \times 5 = 18$
	$4 \times 2 = 6$
	$2 \times 7 = 14$
	$9 \times 2 = 16$

2. Missing data

In general, the data of 95 children were analysed, but in some tasks, there were missing data. These include the following. On the symbolic numerical magnitude comparison task, data of one child were lost due to technical problems. One child switched the keys on the non-symbolic numerical magnitude comparison task, so the child's data of this task were excluded from the analyses. Due to technical problems the data of three children on the multiplication verification task were lost. The non-numerical stroop task was not administered in two boys, because of colour blindness. The teacher did not fill out the BRIEF questionnaire for one absent child. One child was absent during administration of the Tempo Test Arithmetic. One child was absent during administration of the general mathematics achievement test. Two children did not complete the test of intellectual ability.

3. Supplementary results

Accuracy data.

For the stroop tasks and computer tasks a combined score of response time and accuracy was computed by dividing the response time by accuracy. This is legitimated because of high accuracy values (Table A1).

Table A1

Accuracy data of the variables for which combined scores were used.

Variable	<i>M</i>	<i>SD</i>
Symbolic NMP	0.95	0.05
Non-symbolic NMP	0.84	0.10
BC numerical stroop	0.99	0.02
IC numerical stroop	0.95	0.05
BC non-numerical stroop	0.98	0.03
IC non-numerical stroop	0.94	0.06
Addition verification	0.88	0.08
Multiplication verification	0.90	0.09
Motor reaction time	0.98	0.04

Note. NMP = numerical magnitude processing; BC = baseline condition; IC = interference condition.

Sex differences.

To examine potential sex differences in our sample, we administered an independent *t*-test (Table A2). Our results show sex differences on the BRIEF questionnaire, the addition verification task and the general mathematics achievement test.

Table A2

Independent t-tests for sex differences on the administered tasks.

Variable	<i>df</i>	<i>t</i>	<i>p</i>
Symbolic NMP	92	-0.29	0.771
Non-symbolic NMP	92	-1.40	0.164
Numerical stroop	93	-1.27	0.206
Non-numerical stroop	91	-1.11	0.270
BRIEF questionnaire	92	-3.14**	0.002
Tempo Test Arithmetic	92	0.06	0.950
Addition verification	93	2.35*	0.021
Multiplication verification	90	0.89	0.377
General mathematics achievement	92	-2.77**	0.007
Reading	93	-0.29	0.770
Intellectual ability	91	0.21	0.836
Motor reaction time	93	-0.62	0.538

Note. NMP = numerical magnitude processing.

* $p < .05$; ** $p < .01$.

Assumptions regression analysis.

There was no multicollinearity between the independent variables (Table A3), and the assumption of homoscedasticity was met (Figures A1, A2, A3, and A4). The figures show a random array of dots evenly dispersed around zero. They show that at each level of the predictor variables (i.e., symbolic numerical magnitude processing, non-symbolic numerical magnitude processing, numerical stroop task, and non-numerical stroop task), the variance of the residuals is constant. There was no autocorrelation (Durbin-Watson = 1.86; Figure A5) and the errors were normally distributed (Figures A6 and A7). The independent variables included in the regression analysis were linearly related to the dependent variable fact retrieval (Figures A8, A9, A10 and A11).

Table A3

Collinearity between the variables of the regression analysis.

	Pearson correlation	Variance inflation factor
Symbolic NMP and Numerical stroop ^a	0.19	1.04
Symbolic NMP and Non-numerical stroop ^b	0.19	1.04
Non-symbolic NMP and Numerical stroop ^c	0.21	1.05
Non-symbolic NMP and Non-numerical stroop ^d	0.10	1.02

Note. NMP = numerical magnitude processing.

^aControlled for motor reaction time and baseline condition numerical stroop; ^b Controlled for motor reaction time and baseline condition non-numerical stroop; ^c Controlled for motor reaction time and baseline condition numerical stroop; ^d Controlled for motor reaction time and baseline condition non-numerical stroop.

Figure A1

Homoscedasticity test for symbolic numerical magnitude processing.

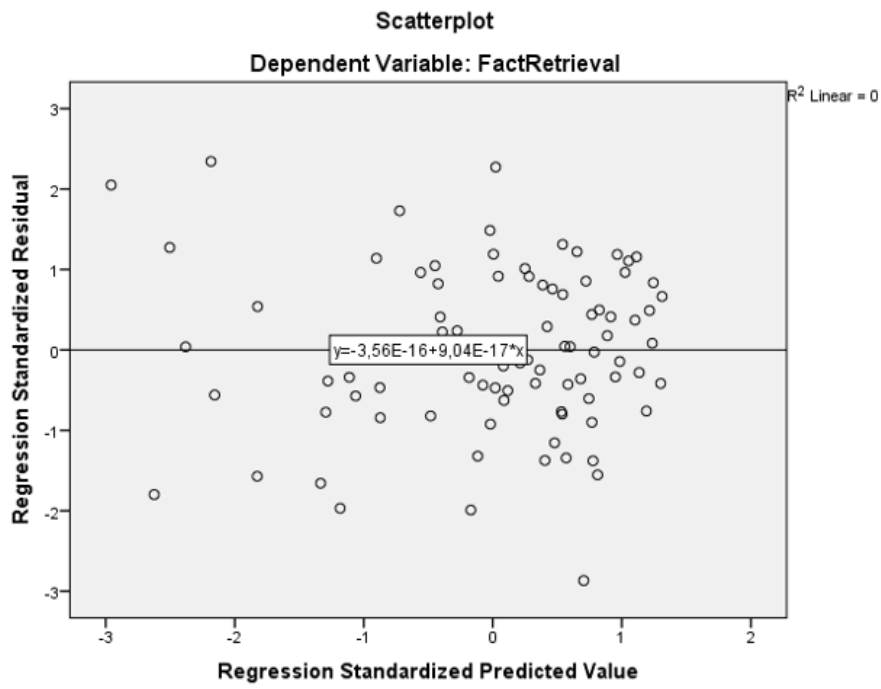


Figure A2

Homoscedasticity test for non-symbolic numerical magnitude processing.

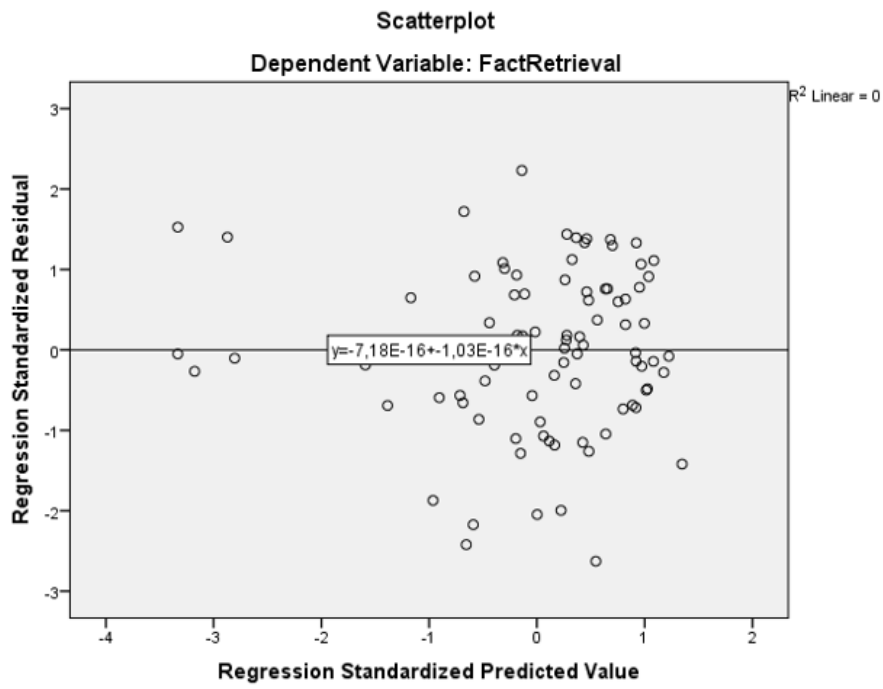


Figure A3

Homoscedasticity test for the numerical stroop task.

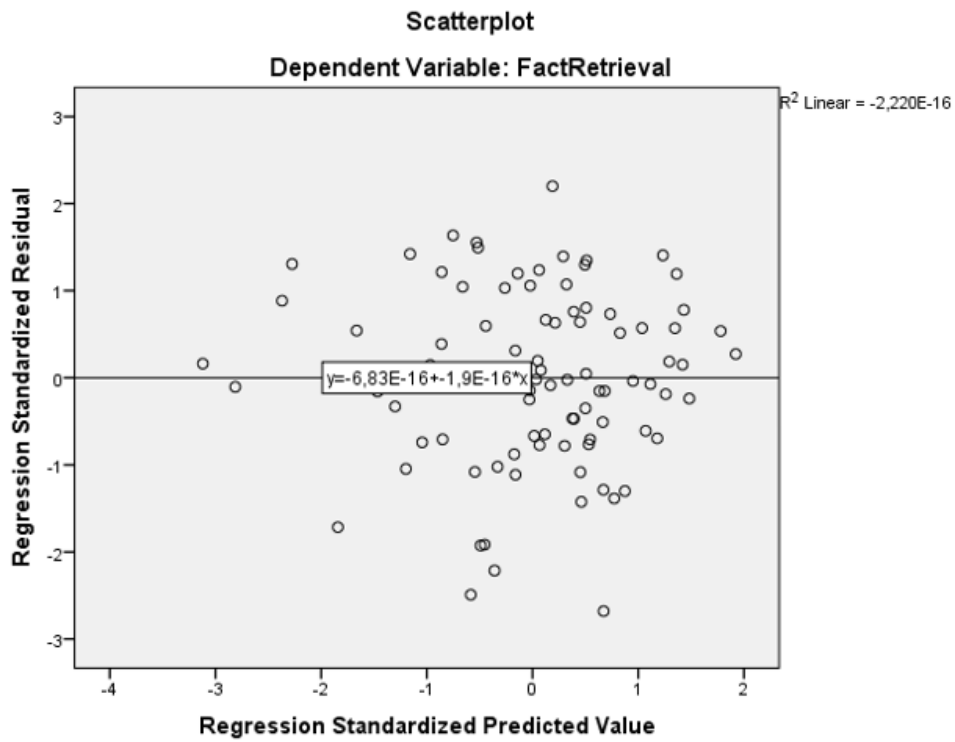


Figure A4

Homoscedasticity test for the non-numerical stroop task.

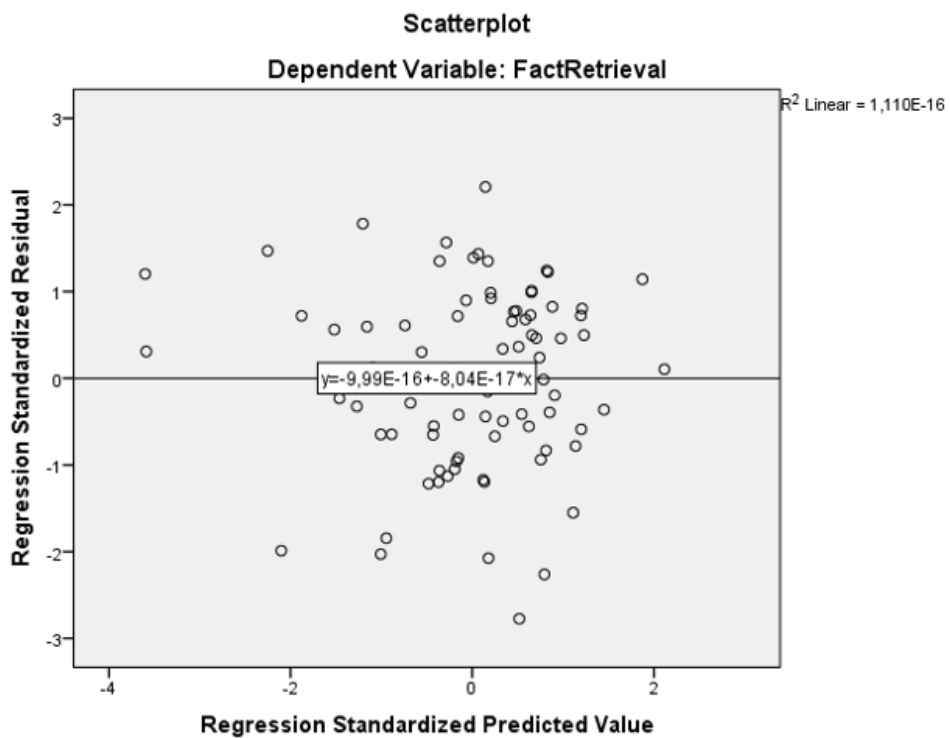


Figure A5

Test for independent errors.

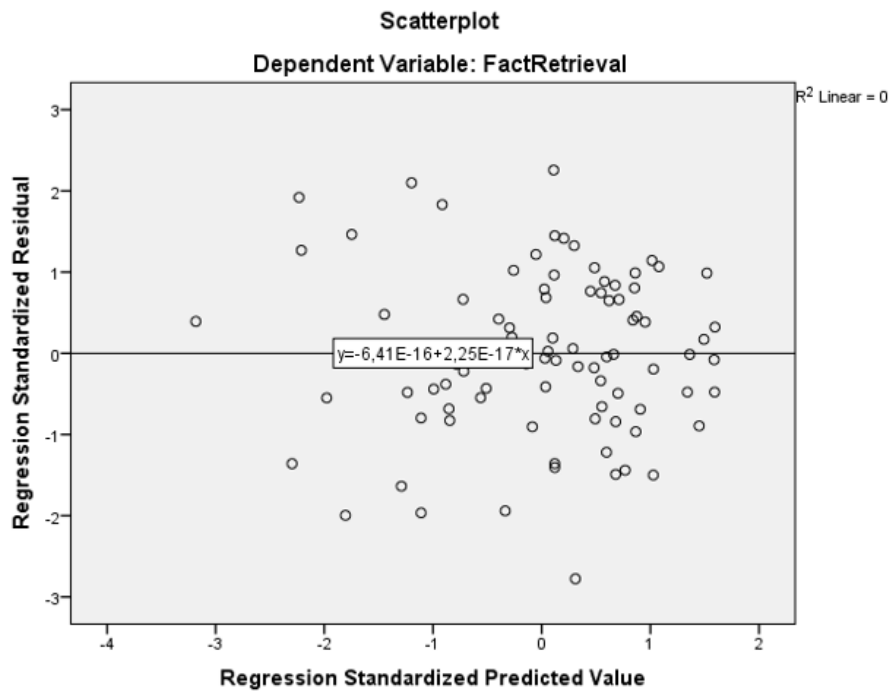


Figure A6

Test for normally distributed errors by means of a histogram.

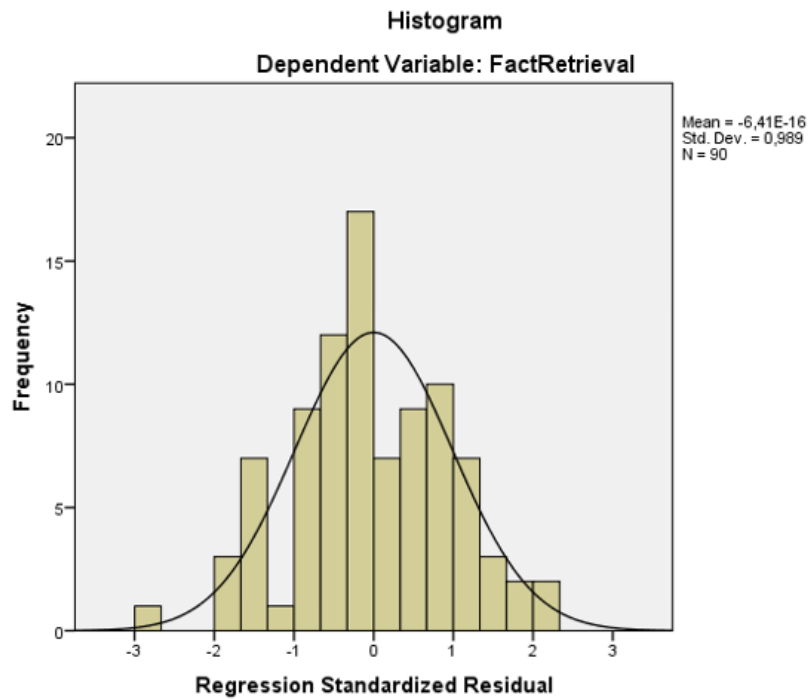


Figure A7

Test for normally distributed errors by means of a P-P Plot.

Normal P-P Plot of Regression Standardized Residual

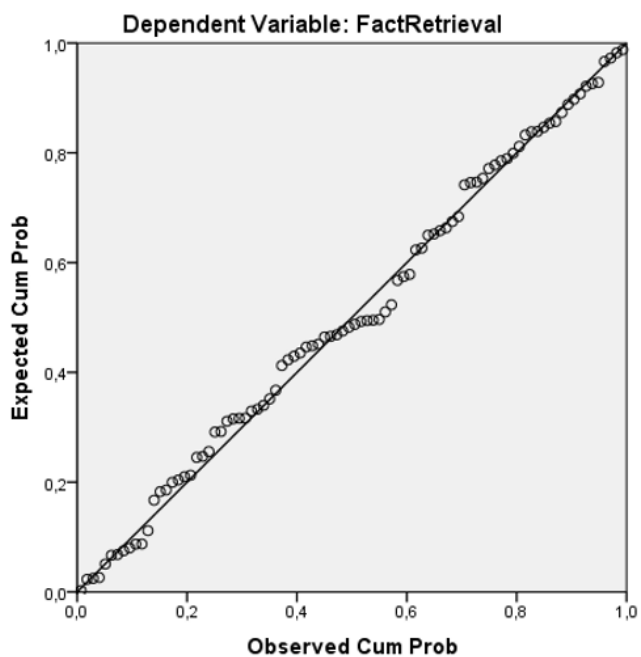


Figure A8

Correlation between fact retrieval and symbolic numerical magnitude processing.

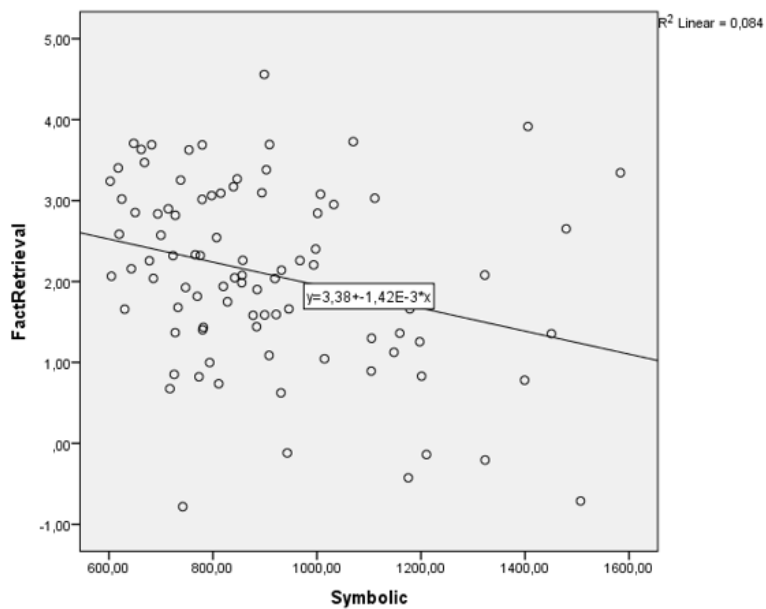


Figure A9

Correlation between fact retrieval and non-symbolic numerical magnitude processing.

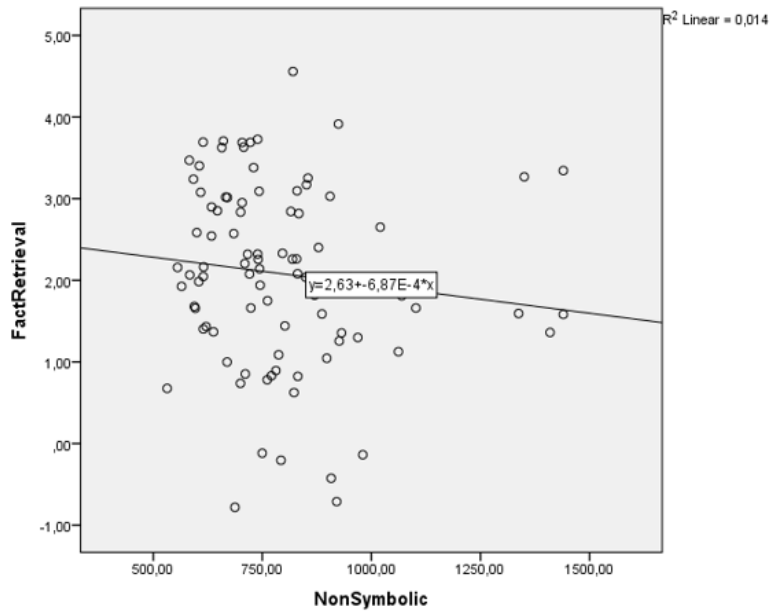


Figure A10

Correlation between fact retrieval and the numerical stroop task.

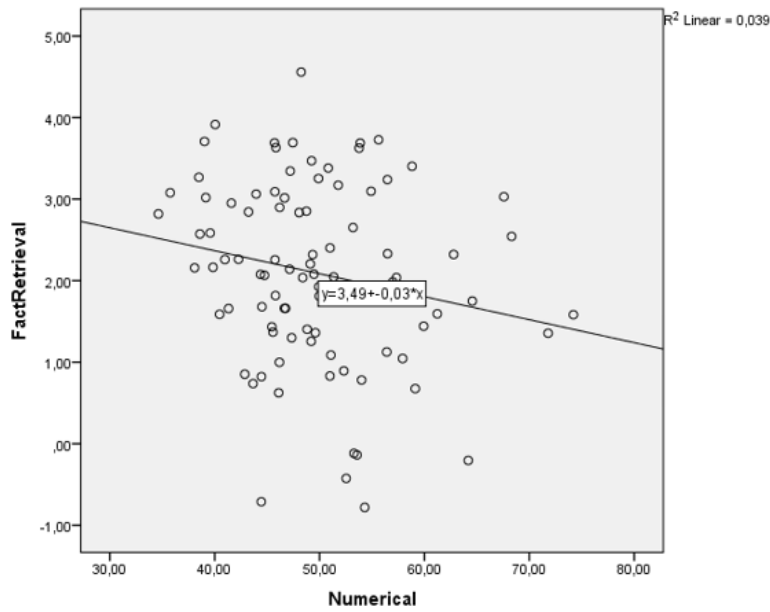
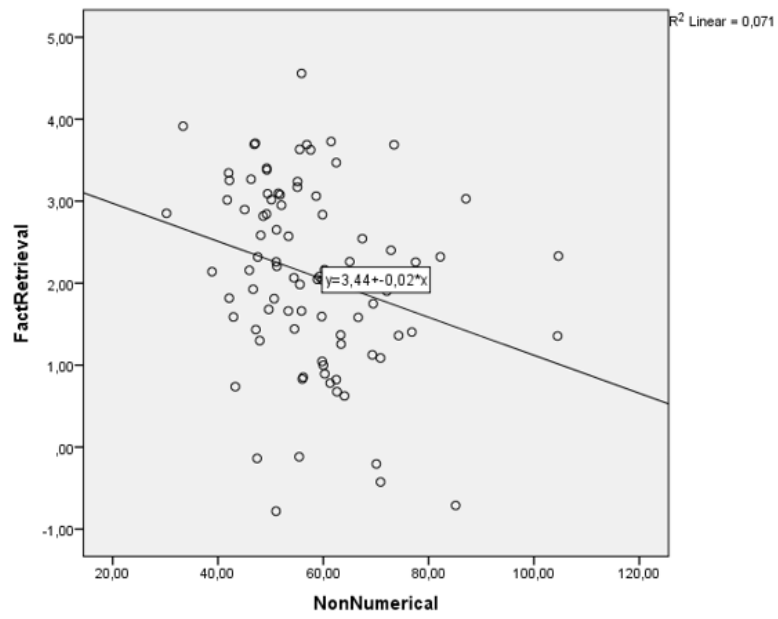


Figure A11

Correlation between fact retrieval and the non-numerical stroop task.



Assumptions ANOVA.

The homogeneity of variance assumption was met for all variables (Table A4), yet not all variables in this study were normally distributed (Table A5). We therefore repeated the analyses with a non-parametric test, i.e. Kruskal-Wallis test. The results of these analyses were entirely similar to those of the ANOVA.

Table A4

Test of homogeneity of variances for the variables of the analysis of variance.

	Levene Statistic	<i>df</i>	<i>p</i>
Symbolic NMP	0.54	2, 87	0.585
Non-symbolic NMP	2.25	2, 87	0.112
BC numerical stroop	1.02	2, 88	0.366
IC numerical stroop	0.43	2, 88	0.652
BC non-numerical stroop	0.02	2, 86	0.981
IC non-numerical stroop	0.15	2, 86	0.863
BRIEF	0.33	2, 88	0.722
General mathematics achievement	0.67	2, 87	0.515
Reading	0.34	2, 88	0.711
Motor reaction time	2.56	2, 88	0.083

Note. NMP = numerical magnitude processing; BC = baseline condition; IC = interference condition.

Table A5

Test of normality for the variables of the analysis of variance.

	Statistic	<i>df</i>	<i>p</i>
Symbolic NMP	0.13	94	0.000
Non-symbolic NMP	0.13	94	0.001
BC numerical stroop	0.09	95	0.037
IC numerical stroop	0.10	95	0.036
BC non-numerical stroop	0.09	93	0.055
IC non-numerical stroop	0.11	93	0.008
BRIEF	0.20	94	0.000
General mathematics achievement	0.13	94	0.001
Reading	0.05	95	0.200
Motor reaction time	0.15	95	0.000

Note. NMP = numerical magnitude processing; BC =baseline condition; IC = interference condition.

Assumptions ANCOVA.

We also tested the two additional assumptions of the analysis of covariance above the assumptions of the analysis of variance. The assumption of independence of the covariate and group was violated (Table A6). However, the assumption of homogeneity of regression slopes was met (Table A7). Since we did not draw conclusions on explained variance or effect size, violation of the assumptions does not influence our interpretations.

Table A6

Dependence of the covariates and cluster condition.

	<i>df</i>	<i>F</i>	<i>p</i>
Symbolic NMP	2, 87	19.25	0.000
Non-symbolic NMP	2, 87	3.61	0.031
Motor reaction time	2, 88	0.07	0.934
BC numerical stroop	2, 88	6.20	0.003
IC numerical stroop	2, 88	4.04	0.021
BC non-numerical stroop	2, 86	0.79	0.458
IC non-numerical stroop	2, 88	3.83	0.025
General mathematics achievement	2, 87	4.82	0.010
Reading	2, 88	3.83	0.025

Note. NMP = numerical magnitude processing; IC = interference condition.

Table A7

Test for homogeneity of regression slopes.

	<i>df</i>	<i>F</i>	<i>p</i>
Symbolic NMP			
Cluster * Motor reaction time	2, 84	1.18	0.312
Cluster * Numerical stroop	2, 84	0.17	0.841
Cluster * Non-numerical	2, 82	0.63	0.536
Cluster * General mathematics achievement	2, 83	0.60	0.553
Cluster * Reading	2, 84	0.94	0.397
Non-symbolic NMP			
Cluster * Motor reaction time	2, 84	1.11	0.335
Cluster * Numerical stroop	2, 84	3.77	0.027
Cluster * Non-numerical stroop	2, 82	2.11	0.127
IC Numerical stroop			
Cluster * Symbolic NMP	2, 84	0.06	0.941
Cluster * Non-symbolic NMP	2, 84	2.52	0.087
Cluster * BC numerical stroop	2, 85	0.19	0.831
IC Non-numerical stroop			
Cluster * Symbolic NMP	2, 82	0.28	0.755
Cluster * Non-symbolic NMP	2, 82	1.39	0.256
Cluster * BC non-numerical stroop	2, 83	0.25	0.777

Note. NMP = numerical magnitude processing; IC = interference condition; BC = baseline condition.

