This is the published version of a paper published in *Frontiers in Psychology*.

**Citation for the original published paper (version of record):**

http://dx.doi.org/10.3389/fpsyg.2016.01062

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

**Permanent link to this version:**
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-124358
Untangling the Contribution of the Subcomponents of Working Memory to Mathematical Proficiency as Measured by the National Tests: A Study among Swedish Third Graders

Carola Wiklund-Hörnqvist1,2*, Bert Jonsson1, Johan Korhonen3, Hanna Eklöf4 and Mikaela Nyroos5

1 Department of Psychology, Umeå University, Umeå, Sweden, 2 Umeå Center for Functional Brain Imaging, Umeå, Sweden, 3 Faculty of Education and Welfare Studies, Åbo Akademi University, Vaasa, Finland, 4 Department of Applied Educational Science, Umeå University, Umeå, Sweden, 5 Department of Education, Umeå University, Umeå, Sweden

The aim with the present study was to examine the relationship between the subcomponents in working memory (WM) and mathematical performance, as measured by the National tests in a sample of 597 Swedish third-grade pupils. In line with compelling evidence of other studies, individual differences in WM capacity significantly predicted mathematical performance. Dividing the sample into four groups, based on their mathematical performance, revealed that mathematical ability can be conceptualized in terms of different WM profiles. Pupils categorized as High-math performers particularly differed from the other three groups in having a significant higher phonological ability. In contrast, pupils categorized as Low-math performers were particularly characterized by having a significant lower visuo-spatial ability. Findings suggest that it is important for educators to recognize and acknowledge individual differences in WM to support mathematical achievement at an individual level.

Keywords: National tests, mathematics, working memory, pupils, education, cognitive profiles

INTRODUCTION

One topic in education that has been receiving rapidly growing attention is the learning of mathematics. In Sweden, mathematics is one of the subjects in school that has the highest failure rate among students. In a related vein, according to the National Center for Educational Statistics (National Center for Education Statistics [NCES], 2011), only 40% of fourth-grade students and 35% of the eight-grade students performed at or above the proficiency level (i.e., solid academic performance for the assessed grade) in math. Likewise, in international comparisons such as Trends in International Mathematics and Science Study (TIMSS, 2007, 2011), Sweden, as well as many other countries, has showed a negative trend for educational achievement in mathematics during the last decade and the Swedish government has allocated a lot of financial resources to find interventions to prevent this trend to continue. Introducing a National test in mathematics in grade three, which is the focus in the current study, is one among other political decisions.

It is widely accepted that there are individual differences in children’s cognitive ability to learn and acquire knowledge for scholastic achievement (see e.g., Engle et al., 1999; Hitch et al., 2001). Working memory (WM) is a cognitive concept thought to play a central role for the development of reading and mathematical skills. Mathematics builds on several cognitive abilities and we know from a wealth of literature that individual differences
in working memory capacity (WMC) are related to mathematical performance and further academic success (Hitch et al., 2001; Reukhala, 2001; Bayliss et al., 2003; Jarvis and Gathercole, 2003; Gathercole et al., 2004; Swanson and Kim, 2007; Bull et al., 2008; Swanson et al., 2008; De Smedt et al., 2009; Gathercole and Dunning, 2010; Menon, 2010; Geary, 2011a,b; Dumontheil and Klingberg, 2012; Nyroos and Wiklund-Hörnqvist, 2012; Frisovanden Bos et al., 2013; Li and Geary, 2013; Bergman-Nutley and Klingberg, 2014), above and beyond measures of socioeconomic status (Dulaney et al., 2015), language skills (Lee et al., 2004), and general intelligence (Alloway and Alloway, 2010). Despite that individual differences in WMC have been shown to influence behavioral measurements in mathematical proficiency, less is known how corresponding differences exist between the subcomponents in WM, as proposed by Baddeley (2000), and the National subtests in mathematics.

Here, we investigate individual differences in WM in relation to performance at the National curriculum tests in mathematics in a sample of Swedish third grade pupils (N = 597). To further investigate if levels of mathematical proficiency can be conceptualized in terms of different WM profiles, we split the sample into four groups derived from overall mathematical performance. The delineation of different WM profiles provides a more complete picture of the influence role of cognitive abilities involved in mathematics. However, it also opens up for an important empirical question: How are the different subcomponents in the original tripartite model of WM (Baddeley, 2000) related to different mathematical proficiency levels?

**Working Memory**

Working memory refers to our ability to temporarily store and manipulate information needed while executing complex cognitive tasks such as numerical and arithmetic processing, problem solving and reasoning (Alloway and Alloway, 2010; Menon, 2010). The most widely used theoretical model for WM is the multicomponent model initially proposed by Baddeley and Hitch (1974) and later revised by Baddeley (2000). According to Baddeley (2000) revised model, WM is composed by four components: the central executive, the phonological loop, the visuo-spatial sketchpad and the episodic buffer (Baddeley, 2000). The central executive is a domain-general attentional control system involved in several processes such as the simultaneously storing and processing of information while handling complex tasks (e.g., mathematics). It is supported by two domain-specific slave systems: the visuo-spatial sketchpad and the phonological loop. The visuo-spatial sketchpad temporarily store visual and spatial information, whereas the phonological loop is involved in the temporarily storage and rehearsal of auditory and phonological information (Baddeley and Hitch, 1974; Baddeley, 2000). Baddeley (2000) added a fourth component, the episodic buffer. The episodic buffer is hypothesized to be responsible for integrating information from the subsystems and long term memory under the supervision of the central executive (Baddeley, 2000). Hitch et al. (2001) investigated WM through complex WM tasks and the possibility of its ability in predicting children’s school performance as measured 1 year later. They found support for the hypothesis that WMC contains a combination of domain-specific and domain-general resources, which in the Bayliss et al. (2003) study has been identified as a general processing component, verbal storage component, and a visuospatial storage component.

**Mathematics**

Mathematics is one of the fundamental skills a child needs to master to successfully progress through the school years (Parsons and Bynner, 2005; Geary et al., 2012). A meta-analysis using six longitudinal datasets found that children with low mathematical ability when entering school/at school entry mostly stayed behind throughout schooling, independent of gender or socio-economic status (Duncan et al., 2007; see also Andersson, 2010; Dulaney et al., 2015 for related findings). Thus, identifying and understanding factors important for mathematical achievement will provide valuable knowledge aiding in the development of appropriate educational methods aimed at enhancing mathematical learning (Turley-Ames and Whitfield, 2003; Gersten et al., 2005; Holmes et al., 2009; Geary, 2011a; Skolverket [National Board of Education], 2011; Dulaney et al., 2015; Ekstam et al., 2015; Swanson, 2015). In a large-scale study, Bryant et al. (2000), compared how teachers rated the behavioral characteristics of 870 pupils identified as having mathematical difficulties with 854 pupils with other educational difficulties. Results showed that the difficulty in carrying out multi-step problems was one feature that specifically differentiated pupils with mathematical difficulties from the other pupils (Bryant et al., 2000), which corresponds to the function of WM.

Mathematics is an umbrella term encompassing a broad variety of competencies, tapping different roles/functions such as switching between operations, strategies, and mental models while solving a task (Mayer, 1998; National Council of Teachers of Mathematics [NCTM], 2000; Niss, 2003; Lépine et al., 2005; Alloway, 2006; Abhakorn, 2008; Litheron, 2008; Litheron et al., 2010; Mee-yin Chan and Suk-han Ho, 2010; Lgr 11, 2011) and known to be closely related to WM (see Peng et al., 2016: for a meta-analysis; Raghubar et al., 2010 for a review). Up to date, the relation between mathematics and WM has mainly been focused on WM as a potential predictor for overall mathematical performance (see Raghubar et al., 2010 for a review; Swanson and Jerman, 2006). For example Hitch et al. (2001) found that the predictability of the complex WM tasks, a year later, accounted for 27% of the variance in basic mathematical skills. However, when considering the variety of mathematical subdomains it seems likely that the contribution of WM and the different subcomponents, within Baddeley (2000) tripartite model of WM, will vary as a function of mathematical domain.

**Working Memory and Mathematics**

The importance of the different subcomponents in the tripartite model of WM (Baddeley and Hitch, 1974; Baddeley, 2000) for mathematical achievement has been investigated using different designs and populations. It is widely accepted that the contribution of executive WM resources for mathematical achievement and performance are crucial, but less consistency
exists about the role of the two slave systems for mathematics. Findings from longitudinal studies with young children are mixed. Using a longitudinal design, De Smedt et al. (2009) suggested that mathematical performance in first grade children are related to individual differences in visuo-spatial ability, but with a shift toward reliance on the phonological ability as a function of increased age (i.e., second grade; De Smedt et al., 2009; see also Hecht et al., 2001). In contrast, other studies have emphasized the visuo-spatial ability as crucial (Bull et al., 2008; Geary, 2011b). For example, in a sample of preschoolers (mean age: 4.6 years) Bull et al. (2008) found in their longitudinal study that visuo-spatial WM, but not phonological WM, predicted mathematical performance at the end of the third grade of primary school.

Similar findings have also been obtained in cross-sectional studies with older children (e.g., Holmes and Adams, 2006). In a sample of typically normal developing 8- and 9-year-old children, Holmes and Adams (2006) found that in both age groups, measurements of the central executive and visuo-spatial ability predicted curriculum-based mathematical performance. Moreover, for older children the measurement capturing the phonological loop predicted performance on easy mathematical tasks, but not performance on difficult tasks (Holmes and Adams, 2006). The latter finding might indicate a gradual shift toward the initial use of cognitive strategies relying on a verbal code (McKenzie et al., 2003), but the efficiency of those strategies might still be in its infancy and related to the degree of task demands (Holmes and Adams, 2006; Raghubar et al., 2010) or moderated by individual differences in WMC (Swanson, 2015).

However, some conflicting evidence exists regarding the role of the subcomponents in WM for mathematical performance related to age. Whereas some studies emphasizes the importance of verbal WM for mathematical performance with increasing age (Swanson and Kim, 2007; De Smedt et al., 2009; Van de Weijer-Bergsma et al., 2015), others found evidence for visuo-spatial WM as important (Gathercole and Pickering, 2000; Reukhala, 2001; Jarvis and Gathercole, 2003; Maybery and Do, 2003; Meyer et al., 2010). Using a cross-sectional design, Van de Weijer-Bergsma et al. (2015) found evidence for both verbal and visuo-spatial WM as equally important for mathematical performance up to grade four, but thereafter verbal WM takes over up to grade six. Thus, Meyer et al. (2010) investigated whether the contribution of the WM subcomponents changed for mathematical achievement in children at age 8- (second grade) as compared to 9-year-old children (third grade). The results showed that both the central executive and the phonological loop predicted mathematical reasoning during the second grade, but not in the third grade. Instead, the visuo-spatial component of WM was predictive for mathematical ability in third graders (Meyer et al., 2010). Correlation studies have found that the relation between the visuo-spatial WM and standardized curriculum mathematical tests persists even in older children, ranging from 7 to 14 years old (Gathercole and Pickering, 2000; Reukhala, 2001; Jarvis and Gathercole, 2003; Maybery and Do, 2003) and Reukhala (2001) found a significant correlation between mathematical performance and visuo-spatial WM even when controlling for verbal WM in a sample of adolescents 15–16 years old. Together, those results indicate the potential role of individual differences in the different subcomponents in WM for academic success across ages, but also the inconsistency findings across studies independent of study design.

The relation between the different subcomponents of WM and mathematical achievements has also been studied among children with mathematical learning difficulties (MD). Meta-analytic findings suggest that children with MD have lower verbal WM and visuo-spatial WM compared to normal achievers (Swanson and Jerman, 2006; Swanson et al., 2009). Furthermore, they found that differences in cognitive functioning between children with MD and normal achievers were primarily related to differences in verbal WM. In contrast, other studies have only found differences in visuo-spatial WM when contrasting children with MD against normal achievers (McLean and Hitch, 1999; Andersson, 2010; Andersson and Östergren, 2012), and against children with reading difficulties (Landerl et al., 2009). Based on the literature it is not clear which subcomponent of WM that is most crucial for mathematics achievement but taken together, findings indicate that individual variation in WM is associated differently depending on at least age (De Smedt et al., 2009), mathematical skills (Swanson and Jerman, 2006), and the specific mathematical domain (Träff, 2013; Kyttälä et al., 2014). It is therefore important to further investigate the different subcomponents of WM in relation to mathematical achievements (Alloway et al., 2004; see also; Bayliss et al., 2005) and in relation to different mathematical domains.

**Aim with the Current Study**

The current study had two goals in mind. The first aim was to investigate the relationship between the WM subcomponents and performance in different mathematics domains (as measured by the National tests in Sweden) in a large and representative sample of Swedish grade 3 pupils. Our secondary aim was to delineate cognitive profiles in relation to mathematics performance by differentiating pupils into mathematical subgroups derived from overall mathematical performance. It was hypothesized that individual differences in WMC would predict mathematical performance overall, but less clear remained how the different subcomponents in WM were predictive for the different subtests in mathematics. Second, we hypothesized that individuals with lower mathematical performance would show a different cognitive profile compared to those performing at a higher level in mathematics. Based on prior studies and the age of the current sample, we expected children with lower mathematical proficiency to be more impaired in both visuo-spatial and phonological WM than children with higher mathematical proficiency.

**MATERIALS AND METHODS**

**Participants**

In the present study, a total of 597 Swedish third grade pupils ($M = 9.34$ years, $SD = 0.30$) participated, 305 girls ($M = 9.35$ years, $SD = 0.30$) and 292 boys ($M = 9.34$ years, $SD = 0.30$).
The sample came from 39 different regular school classes located in five different municipalities. The schools were chosen in order to represent the larger range of geographic and demographic status, based on a grouping by the Swedish Association of Local Authorities. The head teacher and teacher in, respectively, school were contacted and asked to participate. Written informed consent from parents was obtained according with the Declaration of Helsinki, and all children approved to participate. All pupils were assessed at the end of the spring term. The study was approved by the Region Ethical Review Board, Sweden.

Materials
Mathematical Proficiency
Mathematical proficiency was assessed by the Swedish National tests in mathematics for grade 3 pupils. The National tests are state-mandated, curriculum-based tests given in different core subjects to pupils in grade 3, 6, and 9 in compulsory school. The National tests in mathematics for grade 3 pupils consists of seven different subtests covering several different mathematical domains to evaluate a number of syllabus goals (Table 1). The tasks varied in form (from plain numbers to larger tasks) as well as required different methods of expression (e.g., drawing and writing). One subtest was a group assignment and is therefore excluded in the analysis.

As the purpose with the study was to delineate how individual differences in WMC was related to curriculum-based mathematical competencies (see Table 1), we analyzed tests validated and designed to assess those skills according to the National Board of Education in Sweden. The six subtests included in the analysis were: Algorithms and Statistics (maximum score 20, to reach solid academic performance for the assessed grade a score of 14 was required), Fraction (maximum score 13, to reach solid academic performance for the assessed grade a score of 8 was required), Geometry (maximum score 14, to reach solid academic performance for the assessed grade a score of 9 was required), Number understanding and mental arithmetic (maximum score 21, to reach solid academic performance for the assessed grade a score of 14 was required), Problem solving (maximum score 8, to reach solid academic performance for the assessed grade a score of 5 was required) and Time, area and volume (maximum score 13, to reach solid academic performance for the assessed grade a score of 8 was required).

The internal consistency statistic between the subtests was good, with Cronbach’s α = 0.81 (Kline, 2000). The purpose with the National mathematical tests in Sweden are mainly summative but also intended to be used as a formative instrument in which the teacher uses the results as a pupil’s knowledge profile to further support the progress of the pupil’s mathematical proficiency at an individual level.

TABLE 1 | Different syllabus goal in mathematics tested by different subtests (Skolverket [National Board of Education], 2012).

<table>
<thead>
<tr>
<th>The National mathematical subtests</th>
<th>Syllabus goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithms and statistics</td>
<td>The algorithm and statistics assessment tests the pupil’s ability to use and analyze mathematical concepts, use appropriate mathematical methods and to use different mathematical forms of expression to communicate the outcome.</td>
</tr>
<tr>
<td>Fraction</td>
<td>The fraction assessment tests the ability to compare, rank, and divide numbers within integers 0–100 by using different types of illustrations. It also tests the ability to divide wholes into fractions, to describe, compare, identify, and naming fraction as simple fraction</td>
</tr>
<tr>
<td>Geometry</td>
<td>The geometry assessment tests the ability to use rudimentary geometric concepts to describe the characteristics of geometric objects, their position and compare how they relate to other geometric objects by using mathematical forms of expression to communicate.</td>
</tr>
<tr>
<td>Number understanding and Mental arithmetic</td>
<td>The number understanding test the pupils ability to read and write numbers and symbols, and showing digits value in numbers within integers 0–1000. It also assesses pupil’s ability to describe simple number sequences, and managing mathematical equalities within integers 0–20. The mental arithmetic assessment tests the pupil’s ability to solve mathematical problems in the four rules of arithmetic without using written working out as an aid. Numbers and answers are within integers 0–20 also some simple numbers are within an enlarged number area.</td>
</tr>
<tr>
<td>Problem solving</td>
<td>The assessment for problem solving tests the pupil’s ability to formulate and solve problems by using mathematical concepts and mathematical reasoning when communicate and explaining the problem as well as its outcome. The tasks require the pupil to use both symbols and words when calculating, describing and communicate the solution.</td>
</tr>
<tr>
<td>Time, area, and volume</td>
<td>The time, area, and volume assessment tests the ability of measuring simple comparisons, and estimations of different lengths, areas, volumes, and times, using common and appropriate units of measurement to express the results. The task tests the pupil’s ability to initiate and follow a mathematical reasoning.</td>
</tr>
</tbody>
</table>

Operation Span
Measurements with intention to capture individuals’ WMC is often assessed by using complex span tasks, which requires participants to simultaneously process and maintain some information (Conway et al., 2005). The Unsworth et al. (2005) Operation span (Ospan) has shown good internal consistency (0.78) and test–retest reliability (0.83) and is a widely accepted measure of WMC (Turner and Engle, 1989; Klein and Fiss, 1999; Conway et al., 2005; Unsworth et al., 2005; Chein et al., 2011). In the computerized Osplan task the participants are asked to remember a series of letters while performing a concurrent task in which they judge whether a math equation is true or false (for full task descriptions, see Unsworth et al., 2005). In the current study, Osplan was age-adapted such that simpler
mathematical operations were used (i.e., addition, with the sum of integers always in the range of 3–9) but with the same set of high frequency letters in the operation letter strings as in the original version of Ospan (cf. Unsworth et al., 2005). The set-sizes proceeded in fixed level from two sets with no predefined highest level, as long as two consecutive sets at any span length were correctly answered. The dependent variable was the number of correct recalled letters in the correct position (Nyroos et al., 2015).

Digit Span
For the measurement of the phonological WM the digit span (forward and backward) was used (McLean and Hitch, 1999). The computerized digit span was adapted from WISC-IV which has shown to have good internal consistency and re-test reliability ranging from 0.80 to 0.89 (Flanagan and Kaufman, 2004).

A composite of forward and backward digit span (number of correct trials across tasks) served as our measurement for phonological WM (Baddeley, 2000; D’Amico and Guarnera, 2005). In the digit span, numbers ranging from 1 to 9 were displayed on the computer screen at a rate of one number per second. The child is asked to respond by recall the numbers in the correct order (forward and backward, respectively) by pressing the corresponding number on the keyboard. Trials increased from two to a maximum of nine numbers in length, with two trials for each span length. Testing continued until a child failed to repeat two sets at any particular span length. The raw scores for both digit span forward and backward were calculated as a composite of forward and backward digit span (number of correct trials across tasks) served as our measurement for phonological WM (Baddeley, 2000; D’Amico and Guarnera, 2005). Block span measures an individual’s capacity to remember blocks, forward and backward, and is a commonly used measure of visuo-spatial WM (McLean and Hitch, 1999). Spatial span forward is a measure of the visual-spatial storage component of WM (i.e., the visuo-spatial sketchpad) and spatial span backward is a measure of the storage and processing components of visual-spatial WM (i.e., the visuo-spatial sketchpad plus central executive components: Lui and Tannock, 2007). In the block span, 16 green blocks were presented at the computer screen, arranged as a four-by-four square, with one block at a time randomly flashing red at a rate of one box per second. The child is then asked to remember the sequence of blocks displayed red and then respond by recall the sequence on a new square with 16 green blocks, either in the same order as presented (block span forward) or in the opposite order (block span backward). Testing continued until a child failed to repeat two sets at any particular span length, and scores were calculated in the same way as in the digit span task.

Procedures
National Tests in Mathematics
The National tests in mathematics were scheduled to be administered within a 10-week period in the end of the spring term at specific dates decided by the school. The different subtests were conducted by the responsible teacher which all had received specific written instructions and scoring guidelines from The Swedish National Agency for Education (2007; in order to maintain equality in the test procedure as well as at the scoring procedure).

Working Memory Battery
The tasks were administered by two trained research assistants. All tests were assessed individually in front of a computer. The tasks were administered in groups of one to three at the school. Before the session, the participants received information about the confidentiality of individual test results, verbal instructions and each task started with practice trials to ensure they understood the task. To prevent misunderstandings, the pupils were encouraged to ask questions before the assessment took place and written instructions were also provided at the computer screen before each task started during the session. The data collection for the measurements of WM was administered within a 4-month period at the end of the autumn term and ended in the beginning of the spring term just before the period when the National tests in mathematics started.

Statistical Analysis
To investigate how the different WM subcomponents (verbal, visuo-spatial, and central executive) predicted mathematical performance across different mathematical domains, a series of multiple regression analyses were performed. To pursue to what extent mathematical proficiency differed with regard to the specific subcomponents of WM, the sample was divided into quartile groups derived from their overall mathematical performance. Instead of dividing the sample in groups based on cognitive performance, we used the total score of the National tests as cut-off criteria. This resulted in four new groups: Low (n = 144), Low-Average (n = 133), High-Average (n = 165), and High mathematical proficiency group (n = 154). The cut-off score of quartiles was motivated on the basis of using the same procedure as used in prior studies (Jordan et al., 2003; Swanson and Jerman, 2006; Cirino et al., 2015). We used a score at or below the 25th percentile as the cut-off for the group labeled as Low. Given the relatively large sample size in the current study, instead of collapsing children that scored at average (between the 25th and 75th percentile) into one group, we divided those into the Low-Average and the High-Average group. The Low-Average refer to those scoring between the 25th and 50th percentile while the High-Average refers to those scoring between the 50th and the 75th percentile. Finally, the fourth group, labeled as High scored at or above the 75th percentile. Recent research has emphasized that individual differences in cognitive ability varies a lot within the same educational grade (Van de Weijer-Bergsma et al., 2015).
Here, we focused on the whole range by including a large-scale sample of children in regular schooling and within the same educational grade. Multiple group confirmatory factor analysis (CFA) was used to investigate if the WM measures were invariant (i.e., worked similarly well) across the performance groups. This is done by comparing a series of nested models from the least to the most restrictive model. If the more restrictive model does not significantly worsen the fit of the model, measurement invariance is supported. We used the chi-square ($\chi^2$), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA) as overall model fit indices. A non-significant result for the $\chi^2$, values over 0.90 for CFI, and values under 0.08 for the RMSEA indicate good model fit (Marsh et al., 2004). To compare scores was within acceptable normal distribution (Finney and

One participant was recognized as an outlier (mathematical performance below the mean) as between subject factor. Bonferroni correction was applied to correct for multiple testing. The rationale for this analysis was to get a better understanding of how mathematical ability can be understood in terms of WM profiles.

RESULTS

Descriptive results for the overall WM and its subcomponents as well as overall mathematical proficiency and the different mathematical domains are presented in Table 2. One participant was recognized as an outlier (mathematical performance below 2 SDs of the average mean) and therefore excluded from the analysis. A normal distribution analysis showed that the skewness (−1.895) and kurtosis (5.956) for the overall mathematical score was within acceptable normal distribution (Finney and Di Stefano, 2006) but with a tendency for the majority of the pupils to perform well. For the overall WM scores, the analyses of skewness (0.077) and kurtosis (0.046) revealed that the WM scores were normally distributed.

The regression analyses (see Table 3) were conducted using each mathematical subtest as the dependent variable and the three WM measurements as independent variables. In addition, the total mathematical score was also examined within the regression analysis.

As can be seen in Table 3, individual differences in WMC significantly predicted mathematical performance for all National mathematical subtests. Although, the degree of significant contribution for the different subcomponents in WM varied across mathematical subtests (see Table 3).

Next, to identify individual WM profiles, the sample was divided into quartile mathematical groups (Low, Low-Average, High-Average, and High mathematical group) as derived from the general mathematical proficiency score (for descriptive statistics, see Table 2 below). For the four new math groups, there were no significant differences for gender [$\chi^2(3, N = 596) = 1.541$, $p = 0.673$] or chronological age in months [$F(3,592) = 1.94$, $p = 0.31$] between the four groups. To ensure that the WMC measurements were comparable across the four groups, we conducted multiple group CFAs. We used a model that assumed the same factor structure (one overall WM factor, and Ospan, digit span forward and backward, block span forward and backward as factor indicators) across groups but allowed the factor loadings and item intercepts to vary as the comparison model, $\chi^2(20) = 12.35, p = 0.90; CFI = 1.00; RMSEA = 0.00$. We then compared this model with a more restricted model were factor loadings were constrained to equality in all groups but item intercepts were allowed to vary, $\chi^2(35) = 25.596, p = 0.58; CFI = 1.00; RMSEA = 0.00$. Forcing the factor loadings to equality did not significantly worsen the fit of the model, $\Delta\chi^2(15) = 13.246, p = 0.34$. We then compared this model with a fully invariant model were both factor loadings and item intercepts were constrained to equality in all groups, $\chi^2(47) = 39.008, p < 0.05; CFI = 1.00; RMSEA = 0.00$. The

<table>
<thead>
<tr>
<th>Measures</th>
<th>Total sample ($n = 596$)</th>
<th>Low Math ($n = 144$)</th>
<th>Low-Average ($n = 133$)</th>
<th>High-Average ($n = 165$)</th>
<th>High Math ($n = 154$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample ($n = 596$)</td>
<td>17.20 (2.99)</td>
<td>15.34 (2.47)</td>
<td>16.68 (2.96)</td>
<td>17.97 (2.46)</td>
<td>18.57 (3.00)</td>
</tr>
<tr>
<td>Low Math ($n = 144$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operation span</td>
<td>3.02 (1.38)</td>
<td>2.50 (1.33)</td>
<td>2.84 (1.37)</td>
<td>3.27 (1.31)</td>
<td>3.41 (1.34)</td>
</tr>
<tr>
<td>Block span</td>
<td>7.51 (1.56)</td>
<td>6.73 (1.37)</td>
<td>7.36 (1.45)</td>
<td>7.90 (1.48)</td>
<td>7.96 (1.62)</td>
</tr>
<tr>
<td>Digit span</td>
<td>6.66 (1.28)</td>
<td>6.11 (1.09)</td>
<td>6.47 (1.33)</td>
<td>6.80 (1.11)</td>
<td>7.20 (1.33)</td>
</tr>
<tr>
<td>Mathematics (total score)</td>
<td>76.30 (9.41)</td>
<td>63.39 (9.59)</td>
<td>75.33 (1.30)</td>
<td>80.13 (1.41)</td>
<td>85.14 (1.76)</td>
</tr>
<tr>
<td>Algorithm and statistic</td>
<td>16.64 (3.03)</td>
<td>13.19 (2.94)</td>
<td>16.03 (2.10)</td>
<td>17.68 (1.71)</td>
<td>19.28 (0.98)</td>
</tr>
<tr>
<td>Fraction</td>
<td>11.88 (1.51)</td>
<td>10.42 (2.07)</td>
<td>11.95 (0.92)</td>
<td>12.27 (0.88)</td>
<td>12.76 (0.60)</td>
</tr>
<tr>
<td>Geometry</td>
<td>11.18 (1.93)</td>
<td>9.67 (2.18)</td>
<td>10.76 (1.61)</td>
<td>11.54 (1.46)</td>
<td>12.59 (1.04)</td>
</tr>
<tr>
<td>Number understanding and mental arithmetic</td>
<td>19.12 (2.56)</td>
<td>16.42 (3.54)</td>
<td>19.38 (1.72)</td>
<td>20.01 (0.99)</td>
<td>20.47 (0.72)</td>
</tr>
<tr>
<td>Problem solving</td>
<td>6.04 (1.93)</td>
<td>3.94 (2.13)</td>
<td>5.89 (1.44)</td>
<td>7.53 (0.74)</td>
<td>7.53 (0.74)</td>
</tr>
<tr>
<td>Time, area, and volume</td>
<td>11.44 (1.80)</td>
<td>9.75 (2.29)</td>
<td>11.32 (1.39)</td>
<td>12.01 (1.07)</td>
<td>12.51 (0.76)</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01, and ***p < 0.001.
fully invariant model fitted the data as well as the previous model, which clearly indicates measurement invariance across groups, \( \Delta \chi^2(12) = 13.412, p = 0.34 \). In other words, the WMC measures worked in a similar way in all four groups.

The MANOVA with group (Low, Low-Average, High-Average, and High) as the between subject factor and WM (visuospatial, phonological, and executive) as the dependent variables showed a significant group effect for the WM scores, Hotelling’s \( T^2(9,1786) = 14.382, p < 0.001 \). Univariate F-tests revealed significant group effects for all of the WM subtests. Visuospatial ability \( [F(3,592) = 22.31, p < 0.001, \eta_p = 0.102] \), phonological ability \( [F(3,592) = 21.74, p < 0.001, \eta_p = 0.10] \) and for the central executive ability \( [F(3,592) = 14.33, p < 0.001, \eta_p = 0.07] \). For multiple comparisons, Bonferroni adjusted post hoc analyses were performed for each WM subcomponent, and Cohens \( d \) is reported for significant group differences. Cohens \( d \) of 0.2, 0.4, and 0.8 are considered as small, medium and large effect sizes, respectively (Cohen, 1992). Figure 1 (below) depicts the \( z \)-transformed WM scores for each group separately.

**Visuospatial WM**

For visuospatial ability, the Low mathematical group performed significantly lower at the visuospatial tasks when compared to the Low-Average group \( (p = 0.003, d = 0.45) \) and when compared to the other two groups (all \( p’s < 0.001, d = 0.82 \), respectively) indicating an overall lower visuospatial ability in those considered as poor mathematicians. The Low-Average mathematical group had significantly lower scores on the visuospatial task when compared to High-Average \( (p = 0.011, d = 0.37) \) and the High mathematical group \( (p = 0.004, d = 0.39) \) but no significant difference was found between the High-Average group and the High mathematical group \( (p > 0.05) \).

**Phonological WM**

For phonological ability, the Low mathematical group did not differ significantly from the Low-Average group \( (p = 0.08) \), but performed significantly lower when compared to High-Average and the High mathematical group \( (all\ p’s < 0.001, d = 0.63 \) and \( d = 0.90 \), respectively). No significant differences between the Low-Average and High-Average \( (p = 0.129) \) but the High mathematical group performed significantly better at the phonological task when compared to High-Average \( (p = 0.02, d = 0.33) \) indicating that good phonological ability seems to be important for mathematical achievement.

**WM – The Central Executive**

For the complex WM task, no significant differences were found between the Low and the Low-Average group \( (p = 0.20) \)
but the Low mathematical group showed significantly lower performance when compared to both the High-Average and the High mathematical group (all $p's < 0.001$, $d = 0.58$ and $d = 0.68$, respectively). The Low-Average group performed poorer when compared to both the High-Average ($p = 0.039$, $d = 0.32$) and the High mathematical group ($p = 0.002$, $d = 0.42$) but no significant differences between the High-Average and the High mathematical group.

DISCUSSION

The first aim with the study was to investigate the relationship between the subcomponents in WM and performance in different mathematics domains, in a sample of third grade pupils ($N = 596$) in mainstream schools. The predictive role of WM for mathematical performance, as measured by the National tests in Sweden, revealed that individual differences in WMC significantly predicted mathematical performance independent of mathematical domain. This finding confirms prior studies which has suggested a relationship between WM and National curriculum tests (Reukhala, 2001; Jarvis and Gathercole, 2003; Holmes and Adams, 2006; Nyroos and Wiklund-Hörnqvist, 2012; Friso-van den Bos et al., 2013). Our secondary aim was to explore cognitive profiles in relation to mathematics performance by differentiating pupils into mathematical subgroups derived from overall mathematical performance as measured by the National tests in Sweden. Pupils labeled as High mathematical achievers were characterized by having significantly better phonological WM, when compared to the other three groups, whereas those labeled as Low mathematical achievers were characterized by having significantly poorer visuo-spatial WM, as compared to the other three groups. These results are further elaborated below.

Regarding the outcomes from the regression analysis, all three subcomponents in WM (verbal, visuo-spatial, and central executive) significantly predicted overall mathematical performance but the significant contribution varied with respect to mathematical domain. Visuospatial WM, as measured by the Block span task, was the only subcomponent which was a significant predictor across all six mathematical domains. The present findings confirm that individual differences in visuo-spatial ability is crucial for general mathematical proficiency (Gathercole and Pickering, 2000; Reukhala, 2001; Jarvis and Gathercole, 2003; Maybery and Do, 2003; Holmes and Adams, 2006; Meyer et al., 2010; Geary, 2011b). Mathematics as a subject is rather visuo-spatial by nature in which tasks contain diagrams, geometric figures and represent quantities which commonly needs to be mentally visualized to be able to solve a math equation successfully (Li and Geary, 2013). The functional role of the visuo-spatial WM when children solves a mathematical task might be related to the use of mental representations of shapes and/or numbers involved while manipulating mathematical information which clearly put demands on visuo-spatial WM.

Hence, it is worth noting that in the current study the measurement of visuo-spatial WM included both the passive and dynamic aspect of WM (i.e., storage and processing; Meyer et al., 2010) which also might add support for prior findings of the central executive as important for mathematics (Hitch et al., 2001; Bayliss et al., 2003; D’Amico and Guarnera, 2005; Holmes et al., 2008). The predictive value of the visuo-spatial WM found in the current study might be related to the dynamic visuo-spatial WM as the task included both forward and backward block span (see Holmes et al., 2008 for related findings). However, even if there is a differentiation between passive and dynamic visuo-spatial tasks (Raghubar et al., 2010) it seems logic to combine those two when examining its relationship to curriculum-based mathematical tasks, which by nature mostly require some executive resources in terms of simultaneously
maintaining and manipulating information in memory, either by the support of the phonological loop or the visuo-spatial sketchpad.

Surprisingly, the executive part of WM, as measured by a complex span task, was not a significant predictor for fraction. Again, as can be seen in Table 3, the strongest predictor for fraction was related to visuo-spatial WM. In line with prior research, it is plausible to suggest this outcome related to the age of the current sample. The children in the current study were all around the age of nine, and it appears as it is a differentiation approximately around this age in which older children rely more on the phonological loop and children younger than 9 years rely more heavily on the visuo-spatial sketchpad (Kyttilä et al., 2010; Van de Weijer-Bergsma et al., 2015). In that sense, chronological age (age based on the calendar) might be less important and instead favor individual differences in mental age (i.e., age based on intellectual development; Henry, 2001; Henry and MacLean, 2003; Van de Weijer-Bergsma et al., 2015).

For geometry, the only non-significant predictor was phonological WM, which probably is related to the mathematical tasks included. The majority of the tasks in geometry required the pupil to identify and simply describe the characteristics of basic geometric shapes, their position and compare how they relate to other geometric objects. Those cognitive processes might rather rely on automatic retrieval of mathematical facts from long-term memory while maintain and manipulate the visuo-spatial task-specific information in WM, and thereby more executive demanding, without specific demands on the phonological loop (Furst and Hitch, 2000; Reukhala, 2001).

In sum, the results from the regression analysis confirm prior studies by emphasizing individual differences in WMC as predictive for curriculum-based National tests in mathematics. Do note that the skewness and kurtosis for the mathematical proficiency, the results from the mandatory National curriculum batteries, and less commonly done on the basis of educational psychology research domain, this is commonly made on the basis of individuals’ performance derived from cognitive test batteries, and less commonly done on the basis of educational measurements such as the curriculum-based National tests. Therefore, to obtain an ecologically valid profile of mathematical proficiency, the results from the mandatory National curriculum tests in Sweden were used for classification.

An interesting result was the closer inspection of the post hoc classified subgroups in mathematics. When classifying children into low or high performers within the cognitive psychology research domain, this is commonly made on the basis of individuals’ performance derived from cognitive test batteries, and less commonly done on the basis of educational measurements such as the curriculum-based National tests. Therefore, to obtain an ecologically valid profile of mathematical proficiency, the results from the mandatory National curriculum tests in Sweden were used for classification.

Pupils classified as having Low mathematical proficiency performed lower on all WM tasks, but predominantly significant poorer in the task capturing visuo-spatial WM when compared to the other three groups. Those results are important in the light of prior findings from longitudinal studies (e.g., Bull et al., 2008; Geary, 2011b) which has shown that visuo-spatial ability is commonly found as predictive for mathematical ability in children at this age (Bull et al., 2008; Geary, 2011b) and lower visuo-spatial ability has been found as a feature in individuals with mathematical learning difficulties when compared to controls (McLean and Hitch, 1999; D’Amico and Guarnera, 2005; Andersson and Lyytell, 2007; Van der Ven et al., 2013). For example, D’Amico and Guarnera (2005) found evidence for impaired executive and visuo-spatial WM, but not verbal WM in 9-year-old children identified as having poor mathematical ability when compared to normally performing age-matched controls (D’Amico and Guarnera, 2005).

Thus, when delineating the cognitive profile of pupils labeled as High mathematical achievers, another cognitive profile emerged. Compared to the other three mathematical groups, small to large effect sizes between the groups were found (Low, Low-Average, and High-Average, \( d = 0.90 \), \( d = 0.63 \), and \( d = 0.33 \), respectively) indicating that better phonological WM was found among pupils characterized as High mathematical achievers. As indicated by some of the prior research, the relative contribution of the different subcomponents for mathematical performance changes as a function of age. Young children is assumed to rely more on the visuo-spatial component, but as they get older verbal WM gets more involved, thus recruiting the phonological loop. This have mainly been explained in terms of the amplified use of verbal strategies in which children transform numbers and symbols into a verbal code. The trade-off in ages has been suggested to arise around the age of nine, corresponding to the age in the current study.

Notably, performance on the National tests in Sweden is intended to be formative, in the sense that failure to reach set minimum criteria (according to the course syllabi) should alert educators to what mathematical domain which children have difficulties with; subsequently, receiving support in. As indicated by our results, mathematical proficiency could be conceptualized in terms of cognitive profiles indicating cognitive strengths and weakness. Identifying those are of practical significance for educational interventions and methods to further enhance learning by providing a more fine-graded picture of pupils strengths and weakness involved (Bryant et al., 2000; Witt, 2011; Ekstam et al., 2015; Swanson, 2015). Furthermore, we anticipate that this knowledge also will enable the application of appropriate strategies to support children’s learning on an individual level.

Recently, Swanson (2015) investigated the effects of an 8-week strategy intervention among third graders ability to solve problems. The results showed that strategy training had a positive impact on both problem solving and visuo-spatial WM, but the effect of strategy was moderated by individuals WMC (Swanson, 2015). Strategies containing verbal instruction in the absence of visual instruction was only beneficial for those with higher WMC (Swanson, 2015). In contrast, strategy training which contained both verbal and visual instructions produced transfer effects to a task capturing visuo-spatial WM independent of individual differences in WMC. Related to the results in the current study, strategy training might be especially beneficial for those labeled as Low mathematical achievers in the current study as they predominantly were characterized by having a lower visuo-spatial WM. In this respect, it is worth noticing that approximately 10% of pupils in the mainstream classroom are at risk of academic progress difficulties related to WM impairment


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Wiklund-Hörnqvist, Jonsson, Korhonen, Eklöf and Nyroos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.