

## Research Reports

## Creativity as Predictor of Mathematical Abilities in Fourth Graders in Addition to Number Sense and Working Memory

Evelyn H. Kroesbergen\*<sup>a</sup>, Eveline M. Schoevers<sup>a</sup><sup>[a]</sup> Department of Education & Pedagogy, Utrecht University, Utrecht, The Netherlands.

### Abstract

In this study, it was investigated how domain-specific (number sense) and domain-general (working memory, creativity) factors explain the variance in mathematical abilities in primary school children. A total of 166 children aged 8 to 10 years old participated. Several tests to measure math ability, mathematical creativity, number sense, verbal and visual spatial working memory and creativity were administered. Data were analyzed with a series of correlation and regression analyses. Number sense, working memory and creativity were all found to be important predictors of academic and creative mathematical ability. Furthermore, groups with math learning disabilities (MLD) and mathematical giftedness (MG) were compared to a typically developing (TD) group. The results show that the MLD group scored lower on number line estimation and visual spatial working memory than the TD group, while the MG group differed from the TD group on visual spatial working memory and creativity. It is concluded that creativity plays a significant role in mathematics, above working memory and number sense.

**Keywords:** creativity, mathematics, working memory, number sense

Journal of Numerical Cognition, 2017, Vol. 3(2), 417–440, doi:10.5964/jnc.v3i2.63

Received: 2016-06-13. Accepted: 2017-02-11. Published (VoR): 2017-12-22.

Handling Editors: Silke Goebel, University of York, York, United Kingdom; André Knops, Humboldt-Universität Berlin, Berlin, Germany; Hans-Christoph Nuerk, Universität Tübingen, Tübingen, Germany

\*Corresponding author at: Utrecht University, Dept of Education & Pedagogy, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands. E-mail: E.H.Kroesbergen@uu.nl



This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research on individual differences in the development of mathematical cognition has pointed to two main underlying cognitive factors: number sense and working memory (e.g., Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010; Geary, 2010). Number sense (NS) could be seen as the domain-specific precursor of mathematics. Working memory is a domain-general predictor, not only in mathematical development but also in other academic areas. Although several other cognitive factors have been related to math, such as processing speed (e.g., Peterson et al., 2017; Willcutt et al., 2013), or phonological skills (e.g., Barnes et al., 2014; Slot, Van Viersen, De Bree, & Kroesbergen, 2016), working memory and number sense together explain a substantial part of the variance in mathematics (e.g., 48-50% in Kroesbergen & Van Dijk, 2015; 24-28% in Toll, Kroesbergen, & Van Luit, 2016). These two factors are not only related to mathematics in typically developing (TD) children, but they have also repeatedly been found to explain differences between children with mathematical learning disabilities (MLD) and TD children (e.g., Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Mazzocco, Feigenson, & Halberda, 2011; Mussolin, Mejias, & Noël, 2010; Piazza et al., 2010; Raghubar,

Barnes, & Hecht, 2010; Schuchardt, Maehler, & Hasselhorn, 2008). However, from another body of research, on mathematically gifted (MG) children, another domain-general factor appears that may influence mathematics as well: Creativity (e.g., Leikin, Koichu, & Berman, 2009; Mann, 2006; Sriraman, 2005). However, studies investigating the role of creativity in mathematics have taken neither working memory nor number sense into account. The goal of the current study is to compare the role of working memory and number sense with that of creativity in the mathematical abilities of both TD children as well as children with MLD and MG children in fourth grade.

## Number Sense

Number sense (NS) can be defined in very different ways, varying from more restricted definitions as the ability to estimate and understand approximate magnitudes (cf. [Mazzocco et al., 2011](#)) to broader definitions that include counting abilities and even basic math skills (cf. [Jordan, Glutting, & Ramineni, 2010](#)). The definition used in this study is based on the triple-code model of [Dehaene \(1992\)](#). According to this model, number sense is the ability to process numerical information in three different formats: a non-symbolic format (the analogue code), a symbolic-verbal format (verbal code), or a symbolic-visual format (Arabic code). Adequate processing of numbers is logically a basic skill to be able to manipulate numerical information in mathematical tasks. However, the mapping between the non-symbolic and symbolic code seems most essential for further mathematical development ([Mazzocco et al., 2011](#)). For example, [Kolkman, Kroesbergen, and Leseman \(2013\)](#) found that mapping, as measured with comparison tasks and a numberline task, is the best predictor for math skills in Grade 1. [Sasanguie, Göbel, Moll, Smets, and Reynvoet \(2013\)](#) found the same tasks to be the best predictors of mathematics in 6-8 years old children. To understand mathematical tasks, it seems essential to be able to link a numerical symbol to the according numerical meaning. A growing body of research has shown the important predictive value of number sense for later math abilities in longitudinal studies (e.g. [De Smedt, Verschaffel, & Ghesquière, 2009](#); [Friso-Van den Bos, Van Luit, et al., 2015](#); [Fuchs, Geary, Compton, Fuchs, Hamlett, Seethaler, et al., 2010](#); [LeFevre et al., 2010](#); [Kolkman et al., 2013](#)).

The important role of number sense in mathematics is also shown by a large body of research that studied children with MLD. Number sense deficits are found to be a main characteristic of children with serious difficulties in mathematics (e.g., [Mazzocco et al., 2011](#); [Mussolin et al., 2010](#); [Piazza et al., 2010](#)). [Geary et al. \(2007\)](#) made a distinction between children with serious problems or MLD (below 15th percentile) and low-achievers (below 39th percentile). They found that the children with MLD had more serious problems with number sense tasks than the low-achieving children. Following this viewpoint, it could be concluded that part of the MLD children, the ones with the more serious problems, probably have number sense deficits, while the math learning difficulties of the other children with learning difficulties could be attributed to other causes (such as working memory). Other studies, however, found that children with deficits in number sense have comparable math scores as children with deficits in working memory, but that children with deficits in both number sense and working memory have the lowest math scores ([Kroesbergen & Van Dijk, 2015](#); [Toll et al., 2016](#)). This supports the view that number sense deficits can explain MLD, but that other factors are important too.

## Working Memory

Many studies have shown that working memory is an important domain-general predictor of mathematical performance (e.g., Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010; Passolunghi & Lanfranchi, 2012; Raghubar et al., 2010; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). Working memory, or the ability to process and store information on a short term (Baddeley, 2003), is necessary in mathematics, because math often involves storage and processing of numerical information at the same time (Van der Ven, Kroesbergen, et al., 2012). During mathematical problem solving, working memory thus serves as a mental workspace for the – numerical and verbal - information presented in a task, the numbers that have to be manipulated and retrieved numerical facts and procedures. Several longitudinal studies have confirmed that working memory explains a significant part of the variance in mathematics (see Raghubar et al., 2010, for an overview). Even a causal role has been suggested by experimental studies in which children's working memory skills were manipulated (e.g., Bergman-Nutley & Klingberg, 2014; Holmes & Gathercole, 2014; Kroesbergen, Van 't Noordende, & Kolkman, 2014; Söderqvist & Bergman-Nutley, 2015), although the meta-analysis of Melby-Lervåg and Hulme (2013) shows that there is no clear evidence of generalization of working memory training to math skills. It should be noted that in the literature working memory sometimes refers to the entire processing and storage unit as described by Baddeley and Hitch (1974), while other use it to refer only to the central executive, or even more specifically to the executive function of updating (see also Friso-Van den Bos et al., 2013). In this article we refer to the central executive component of the working memory model, which of course also relies on the slave components and has elements of inhibition, shifting and updating in it.

A distinction can be made between verbal working memory and visual spatial working memory skills, for the processing of verbal and visual spatial information respectively. Some authors have found that the visual spatial working memory is most strongly related to mathematics (Schuchardt et al., 2008), while others show that verbal working memory is more important (Friso-Van den Bos, Van der Ven, Kroesbergen, & Van Luit, 2013). Two recent studies seem to suggest that during elementary school the role of visual spatial working memory diminishes while the role of verbal working memory inclines (Van der Ven, Van der Maas, Straatemeier, & Jansen, 2013; Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2015). Early mathematics, in which still many visual materials and models are used, and which requires insight in representations of numerical information, may rely more on visual spatial working memory, but when children grow older, start memorizing basic mathematical knowledge and procedures, and problems are presented in verbal formats, the role of verbal working memory may become more important.

The role of working memory in mathematics is also shown by studies investigating children with MLD. Many studies have shown that children with MLD in general have lower working memory skills than TD children (e.g., Geary, Bailey, & Hoard, 2009; Geary et al., 2007; Raghubar et al., 2010; Schuchardt et al., 2008). It has also been found that children with low working memory skills have low mathematical abilities (e.g., Alloway, Gathercole, Kirkwood, & Elliott, 2008; Kroesbergen & Van Dijk, 2015). Especially visual spatial working memory seems to be impaired in children with MLD (Ashkenazi, Rosenberg-Lee, Metcalfe, Swigart, & Menon, 2013; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013). The study by Ashkenazi and colleagues (2013) showed that MLD children do not use visual spatial working memory resources appropriately during arithmetic problem solving. This is especially relevant because visual spatial working memory seems to be most involved in new math tasks. Thus, limitations in visual spatial working memory may hinder further mathematical development.

On the other hand, studies investigating working memory in gifted students, indeed found higher working memory skills in this group (Dark & Benbow, 1991; Geake, 2009).

## Creativity

How creativity comes into play in the process of mathematical development, is yet largely unknown. Creativity can be defined as the production of novel and useful products (broadly interpreted) within a social context (Plucker, Beghetto, & Dow, 2004). In education creativity is relative; a creation of a product (e.g. solution or idea) is seen as creative when it is novel and useful for a specific student (Leikin, 2009). Creative skills are thought to be beneficial for learning in general (Kaufman & Sternberg, 2010), and learning mathematics in specific (Kattou, Kontoyianni, Pitta-Pantazi, & Christou, 2013), although conclusive evidence is lacking. At least some degree of creativity is required when a child is faced with a problem for which he has no practiced or learned solution (Leikin et al., 2009). Since solving problems is a major feature of mathematics, the role of creativity in mathematics could not be denied. It should be noticed that many of the mathematical tasks children are faced with in primary schools, cannot be qualified as *problem solving*, but rather as applying routine or familiar procedures or even as giving the automatized answer (e.g. in learning multiplication tables). However, previous research has shown that the phase of problem solving is also necessary in the process of learning routine procedure or rote memorization, such as learning the multiplication facts (Van der Ven, Boom, Kroesbergen, & Leseman, 2012), which suggests that creative children may also be better at acquiring basic facts or procedures.

Hadamard (1996) argues that invention or discovery, takes place by combining ideas, whether it is in mathematics or anywhere else. The invention or discovery of solutions for mathematical problems takes place by combining ideas, or ideas and formerly learned knowledge or procedures. Creativity in mathematics thus requires the activation of multiple ideas, and sources of information simultaneously, to think of alternative solutions for mathematical problems. However, only few are useful, fruitful or exceptional. But one needs to construct all possible combinations to find these (divergent thinking), and once having divergent ideas, one can choose and elaborate the most promising one to get to the optimal solutions (convergent thinking; Hadamard, 1996; Koichu, Berman, & Moore, 2006). Experienced problem solvers show different phases in their solution process (Carlson & Bloom, 2005), in which both divergent and convergent thinking play a role. Generally, the focus in math education is on convergent thinking (giving the correct answer). However, excellent mathematics learners are characterized by the understanding that more than one approach can lead to equivalent results and the ability to solve problems in different ways, in other words, are characterized by mathematical creativity (Leikin & Lev, 2007).

The question is what is necessary to become excellent in mathematics. Gifted children have been found to differ from non-gifted children in the ability to find multiple solutions (Leikin & Lev, 2007). They also have on average higher working memory capacity (Geake, 2009) than their non-gifted peers and are therefore supposed to be more able to make combinations between pieces of knowledge and information and to consider more outcomes concurrently. Fewer ideas have to be discounted from the outset in the decision-making process, allowing higher activation of multiple ideas, and sources of information simultaneously, to think of alternative solutions for mathematical problems. A plausible hypothesis is thus that creativity is a crucial factor for excellence in mathematics, above working memory and domain-specific knowledge. Logically, this effect will be larger in mathematical tasks that require at least a certain level of creativity, as in problem solving, than in,

for example, speeded arithmetic tests. It should be noted that in creative thinking, domain general factors such as working memory and executive functions are also needed (see [Silvia, 2015](#) for a review). In studying creativity, it is thus necessary to take domain general factors also into account.

## The Current Study

The goal of the current study was to compare the role of creativity to that of working memory and number sense in mathematical learning. It is the first study to combine these three predictors of mathematics into one model. This makes it possible to study these predictors not only in relation to mathematics, but also to each other. Specific attention is given to children with MLD and MG children, because the profiles of these groups can inform us about what skills are necessary to become proficient or even excellent in mathematics. Three research questions have been formulated:

1. How are different aspects of number sense and working memory related to mathematics in fourth graders? This first question will be answered to investigate the specific contributions of number sense and working memory to math in children from fourth grade. For number sense, both non-symbolic and symbolic tasks will be used, and for working memory verbal and visual spatial tasks. It is expected that the symbolic number sense skills are more strongly related to math than the non-symbolic number sense skills (cf. [De Smedt et al., 2009](#); [Friso-Van den Bos, Van Luit, et al., 2015](#); [Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010](#); [Kolkman et al., 2013](#); [LeFevre et al., 2010](#)). Furthermore, it is expected that verbal working memory is more strongly related to math than visual spatial working memory, because of the age of the children (cf. [Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2015](#); [Van der Ven et al., 2013](#)).
2. Is creativity related to academic math tasks and to creative math tasks? The relations between a general creativity task and two types of math tasks are investigated. The first task reflects the current mathematics curriculum, the second task is designed to measure mathematical creativity, because it asks for multiple solutions. Mathematical creativity requires both creative skills and mathematical skills and can be regarded as a subcomponent of mathematics ([Kattou et al., 2013](#); [Schoevers, Kroesbergen, & Kattou, 2017](#)). It is expected that creativity is related to both math tasks, but the strongest to the mathematical creativity task (cf. [Leikin et al., 2009](#)). Furthermore, it is expected that creativity explains variance in math performance, above the variance explained by working memory and number sense.
3. What are the differences in number sense, working memory and creativity between students with mathematical difficulties and typically and mathematically gifted students? To get more insight into the predictors that may be responsible for mathematical difficulties on one hand and mathematical strengths on the other, the weakest and highest performing children of the sample have been compared to the average performing children. It is expected that children with MLD show deficiencies in number sense and visual spatial working memory (cf. [Kroesbergen & Van Dijk, 2015](#)), and that the MG children are characterized by high working memory, and creativity (cf. [Leikin & Lev, 2007](#)).

## Methods

### Participants

Children were recruited from 11 schools in a large city in the middle of the Netherlands. All children from fourth grade were asked to participate ( $N = 337$ ). The parents of 270 children gave permission by active informed consent. However, results only of the math test were available from 172 children. Children with other disabilities than dyscalculia (e.g., attention deficits or dyslexia) were removed from the data ( $N = 6$ ). This resulted in a sample of 166 children, with a mean age of 9.66 years ( $SD = 0.58$ ), and of which 79 were boys (47.59%). Children who scored below the 20<sup>th</sup> percentile (score V) on the CITO math test, and had a history of low performance in mathematics according to school data ( $>1 SD$  below the group mean) were categorized as MLD. Children with a score above the 80<sup>th</sup> percentile (score I) and a history of high mathematical performance ( $> 1 SD$  above the group mean) were categorized as MG. The remaining children were categorized as Typically Developing (TD; see also Table 1).

Table 1

*Descriptive Statistics of the Main Variables of the Study*

Statistic	TD	MLD	MG
<i>N</i>	130	16	20
% boys	47.70	25.00	65.00
<i>M</i> age (years)	9.64	9.89	9.50
<i>SD</i> age	0.62	0.62	0.43

*Note.* TD = Typically Developing; MLD = Math Learning Disabilities; MG = Mathematically Gifted.

### Instruments

A series of tests was administered to measure students' number sense, working memory, intelligence, creativity, and mathematical ability.

#### Number Sense

A newly developed computer-based number sense test was used (Friso-Van den Bos, Schoevers, Slot, & Kroesbergen, 2015), containing three different number sense tasks, that together adequately reflect the different aspects of number sense, both symbolic and non-symbolic. These tests are based on Kolkman and colleagues (2013) and Friso-Van den Bos, Kroesbergen, and Van Luit (2014). All three tasks were administered on a laptop computer using E-prime 1.2 software (Schneider, Eschman, & Zuccolotto, 2002).

**Symbolic comparison** — In the symbolic comparison task, children were confronted with two numbers below 100 on the screen, and they have to respond which one is the largest number by pressing an 'a' for the left number and an 'l' for the right number. Children were asked to respond as fast as possible. The largest number was equally presented on the left and the right. The task started with a training block of 6 items, which were not used in the scoring. After the training block 33 fixed items were randomly presented to the children. These items differed in the degree of difficulty. The difficult items included two numbers that were close to each other (ratio 0.88; e.g. 14 vs.16). Average and easy items consisted of numbers that had respectively a ratio of 0.75

(e.g. 12 vs.16) and 0.625 (e.g. 10 vs.16). Both accuracy (percentage correct) and reaction time were scored. The internal consistency of the task is high ( $\alpha = .84$ ; Kline, 1999).

**Non-symbolic comparison** — The non-symbolic comparison task is comparable to the symbolic comparison task, with sets of dots instead of numbers. The amount of dots range from 1 – 100. Children were asked to respond as fast as possible but accurately. The largest size of dots was equally presented on the left and the right. The task started with a training block of 6 items, which were not used in the scoring. After the training block 43 items were presented to the children in three different blocks. After each block (of respectively 14, 14 and 15 items) there was a short break. In each block the items were randomly presented. In all items the size or the convex hull (i.e. the smallest measured surface in which the dots can fit) of the dots were manipulated. In 14 items the size of the dots and the convex hull of the dots were respectively incongruent and congruent with the amount of dots. In another 14 items this was the other way around. Furthermore in 15 items the convex hull stayed the same and only the size of the dots was manipulated (incongruent or congruent). The items within each different manipulation also differed in difficulty. Three different ratios were used (0.625, 0.75 & 0.88). An example of one item can be found in Figure 1. Both accuracy and reaction time were scored. The internal consistency of the task is high ( $\alpha = .93$ ; Kline, 1999).

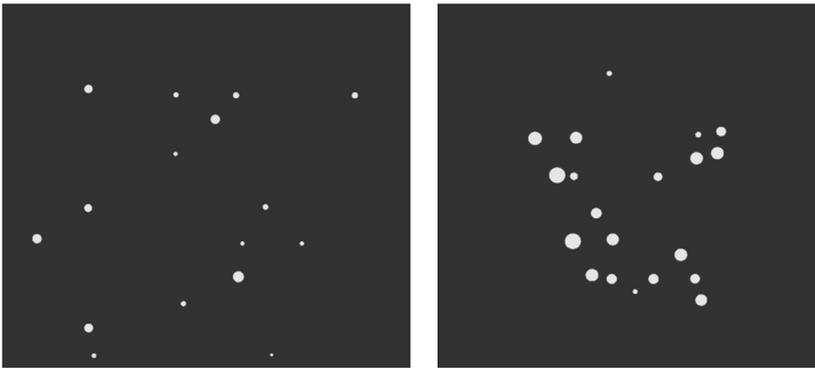


Figure 1. Item of the non-symbolic comparison task. The left group of dots has 16 dots. The right group of dots has 20 dots. The convex hull of the right group of dots is incongruent with the number of dots, but congruent with regard to the size of the dots.

**Numberline task** — The number line estimation task demanded children to place a lever on a number line from 0 to 100 to position the digit that was presented. The test started with two training items in which the child had to place 0 and 100 on the number line. After the training block, 30 test items were randomly presented to the children. The proportion of explained variance ( $R^2$ ) was computed by fitting the answers of each child on a linear curve (see also Kolkman et al., 2013). Internal consistency of the tests was high ( $\alpha = .80$ ; Kline, 1999).

### Working Memory

Two online computerized working memory tasks were administered (Van de Weijer-Bergsma, Kroesbergen, Jolani, & Van Luit, 2016), one for measuring visual spatial working memory and one for verbal working memory.

**Visual spatial working memory** — The Lion game is a visual spatial complex span task, in which children have to search for colored lions (Van de Weijer-Bergsma, Kroesbergen, Prast, & Van Luit, 2015). Children were presented with a 4 x 4 matrix containing 16 cells. In each trial, eight lions of different colors (red, blue, green, yellow, purple) were consecutively presented at different locations in the matrix. Children had to remember the

last location where a lion of a certain color had appeared and they had to use the mouse button to click on that location after the sequence has ended. The task started with two practice trials in which they had to remember the location of the last red and the last blue lion. Children received feedback on their performance after each trial. After the trials, the assessment started which consisted of 20 items in five difficulty levels, in which working memory load is increased by the number of colors that have to be remembered. At Level 1, children had to remember the location of the last red lion. At Level 2, children had to remember the locations of the last red and the last blue lion, and so on (Level 3: red, blue, and yellow; Level 4: red, blue, yellow, and green; Level 5: red, blue, yellow, green, and purple). Items were constructed using randomization with regard to sequence of location and color, with one constraint: items never end with a red lion, since the first response requires the location of the last red lion. The proportion of items recalled in the correct location was used as a measure. The Lion game has excellent internal consistency (Cronbach's  $\alpha$  between .86 and .90), satisfactory test-retest reliability ( $\rho = .71$ ), and good concurrent and predictive validity (Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2015).

**Verbal working memory** — The Monkey game is a verbal span-backwards task, in which children have to remember and recall different words backward. Children heard spoken one-syllable words and had to recall them backwards, by clicking on the written words presented visually in a 3 x 3 matrix. Before the task started, children were presented with four practice sets. In the first two practice sets, children were asked to recall two words forwards. After those sets, children were informed about the backward recall procedure and were presented with two more practice sets in a backward fashion. After each practice children received feedback on their performance. After the training the assessment started, which consisted of 20 items in five difficulty levels (ranging from two words in Level 1 to six words in Level 5) in which working memory load is manipulated by the number of words children have to remember and recall backward. Levels were presented in increasing order. The item sets within each level were constructed using randomization with regard to the sequence of words. The proportion of items recalled in the correct location was used as measure. The Monkey game has excellent internal consistency (Cronbach's  $\alpha$  between .78 and .89) and shows good concurrent and predictive validity (Van de Weijer-Bergsma et al., 2016).

### **Creativity**

The Test of Creative Thinking-Drawing Production (TCT-DP; Urban & Jellen, 1996) was used to measure creativity. The TCT-DP form A was used and took 15 minutes. This test mirrors a holistic concept of creativity (Urban, 2005). Students had to complete a drawing while using some given figural fragments. The TCT-DP was scored according to the guidelines in the manual. Scores were obtained for 14 categories: 'continuations', 'completions', 'new elements', 'connections made with a line', 'connections made to produce a theme', 'boundary breaking that is fragment dependent', 'boundary breaking that is fragment independent', 'humor and affectivity', four categories of 'unconventionality', and 'time'. A total score (max. 72) was obtained by adding the scores of the 14 categories, as recommended by Urban (2005). Two raters scored all tests, the intra-class correlation of the total scores (ICC; two-way mixed consistency single-measures ICC) was high in this study ( $\alpha = .93$ ; Hallgren, 2012).

## Mathematics

Two tests were used to measure mathematical achievement, a standard test that covers a broad range of mathematical tasks and that is typical for current academic math tests, and a test that measures mathematical creativity by requiring multiple solutions.

**Academic mathematical achievement** — Scores from the national Cito mathematics test (Janssen, Scheltens, & Kraemer, 2007), developed by the national institute of educational testing, were used as a measure for mathematical achievement. Different, age-appropriate, versions of the test are administered twice a year by the teacher. The results of the test of middle grade 4 were used for this study. Furthermore, the schools were asked to share the scores of the children from former grades. The test consists of primarily word problems that cover a wide range of mathematics domains such arithmetical operations, geometry, measurement, time, and proportions. Test scores are converted into normed ‘ability scores’, provided by the publisher. The national mean ability score of Dutch fourth grade students is 86 (HCO, 2016). The Cito mathematics test has been shown to be highly reliable; the reliability coefficients of different versions range from .91 to .97 (Janssen, Verhelst, Engelen, & Scheltens, 2010).

**Creative mathematical achievement** — The Mathematical creativity test (MCT), developed by Kattou and colleagues (2013) and translated by Schoevers and colleagues (2017), was used to measure mathematical creativity. The MCT consists of five open-ended questions that could have multiple solutions. The maximum administration time was 45 minutes. Students were required to provide multiple solutions, which were distinct from each other and which were original. Scores were obtained for fluency, flexibility and originality for each question. For the fluency score the number of correct solutions was counted, and divided by the maximum number of correct mathematical solutions provided by one of the students from the sample. To calculate the flexibility score, the number of different types or categories of correct solutions was counted per question and divided by the maximum number of different types of solutions provided by a student in the sample of the study. Originality was calculated by comparing a student’s solution with the solutions provided by all students who participated in the study. The originality score, ranging between 0-1, was given for each question. The rarest correct solution received the highest score: a student was given the score 1 for originality if one or more of his/her answers appeared in less than 1% of the sample’s answers. A score of 0.8, 0.6, 0.4 or 0.2 was given if the frequency of one of more of his/her answers appeared respectively in between 1% and 5%, 6% and 10%, 11% and 20%, and more than 20% of the sample’s answers. Fluency, flexibility and originality scores of each question were added and in this way scores were obtained for the scale of fluency, flexibility and originality. The internal consistency between the three scales was high in this study ( $\alpha = .88$ ; Kline, 1999). The fluency, flexibility and originality scores were transformed to z-scores and averaged to get one measure of mathematical creativity. Interrater reliability was calculated for the scales of fluency and flexibility and was based on a sample of 10 tests scored by three raters. The interrater reliability was high for the scale of fluency (ICC = .96; Hallgren, 2012) and flexibility ( $\kappa = .96$ ; Landis & Koch, 1977). The remaining tests were scored by one of these three raters.

## Intelligence

IQ was estimated based on two subtests of the ‘Nederlandse Intelligentietest voor Onderwijsniveau’ (Dutch Intelligence test for Education level; Van Dijk & Tellegen, 2004), that measure verbal and visual spatial reasoning. In the subtest ‘Categories’, children have to find a logical relation between sets of two words. The

words can be identical, contrary, a sort of, a part from or a cause and meaning to each other. In the subtest 'Foldouts' children get a piece of paper with eight items. The items consist of one three-dimensional figure and five two-dimensional foldouts. One of the five options can be fold in the three-dimensional figure. Children have to choose which one is the right option. The reliabilities of these subtests are good ( $\alpha = .82-.86$ ; Van Dijk & Tellegen, 2004). The standardized scores of the two subtests were averaged, since the two subtests were highly related ( $r = .73$ ).

## Procedure

Data were collected in the fall of 2015 by research assistants with at least a bachelor's degree in Special Education, supervised by the first author. Active informed consent was obtained from the parents of all children involved. Information about students' age and gender was provided by the teacher. The MCT and TCT-DP were administered groupwise by two research assistants in a session of approximately one hour. The IQ tests were administered groupwise in a second 30-minute session within one week of the first session. Testing was standardized by means of elaborate written instruction for the testers. Instructions were read aloud. Additional instruction was provided for individual students when needed. Students were not allowed to copy the work of their fellow-students and to talk during test sessions. After the groupwise administration, children completed respectively the Monkey and Lion Game and the Number Sense tests individually on a laptop computer, supervised by one of the research assistants. The study was approved by the ethics committee of the faculty of Social and Behavioural Sciences of Utrecht University (FETC14-005).

## Data Analysis

The dataset was not complete. Especially in the computerized number sense tests, a relative large amount of the data was missing (17%), mostly due to computer problems. There were also data missing in the creativity test (9.6%), because it was impossible to score some tests. This was for example because children erased many lines and it was difficult to see which ones were erased or not, or they used a very thick pencil, or had very bad handwriting. Missing data were handled by using pairwise deletion. Pairwise deletion of the missing data was not considered problematic since the data in this study were missing in a random pattern (Little's MCAR test (Little, 1988):  $\chi^2(60) = 53.59$ ,  $p = 0.71$ ); pairwise deletion produces consistent estimates of the parameters when data are missing completely at random (MCAR; Allison, 2009).

To examine the relations between mathematics, working memory, number sense, and creativity, Bayesian correlation analyses were conducted, with the statistical package JASP (JASP Team, 2016). Bayesian model selection, using the BIEMS software package (Mulder, Hoijtink, & De Leeuw, 2012), was used to compare regression models and to confirm the differences between the MLD, TD, and MG group (for a detailed introduction see Hoijtink, 2012). IQ was included in all analyses to control for possible general ability effects. Bayesian model selection offers the possibility of using prior knowledge to formulate and evaluate informative hypotheses and to compare competing hypotheses. In addition, this type of analysis enables the testing of multiple hypotheses (as is the case in this study) without the loss of power (due to for example Bonferroni-type corrections, as elaborated in Hoijtink, Klugkist, & Boelen, 2008; Van de Schoot et al., 2011), and is especially suitable for small sample sizes, as it is not based on normality or asymptotic assumptions (Gill, 2008).

For each analysis, specific hypotheses were formulated and tested. Bayesian analyses involve the calculation of the Bayes Factor (or BF) of an informative hypothesis versus the alternative hypothesis. The BF represents

the amount of support from the data in favor of one hypothesis compared to another hypothesis. In addition to the BF, posterior model probabilities (PMPs) can be computed, representing the relative support for a specific hypothesis within a set of hypotheses. If the PMP of one hypothesis is larger than the PMP of the unconstrained hypothesis, the constraints used to specify the hypothesis are supported by the data; and if the PMP of a first hypothesis is larger than the PMP of a second hypothesis, the support in the data is larger for the first than for the second hypothesis. Note that the sum of the PMPs for a set of competitive hypotheses is always one. Moreover, in this type of analysis it is not necessary to provide additional estimates of effect size. The effect size is incorporated into the BF in the sense that a larger effect size results in a larger BF (Kluytmans, Van De Schoot, Mulder, & Hoijtink, 2012). Kass and Raftery (1995) provided more information about cut-offs that help when interpreting generated BFs. A BF below 1 indicates that there is more support for the alternative hypothesis, a BF between 1 and 3 represents a small effect in favor of the informative hypothesis, between 3 and 10 means that there is substantial evidence supporting the informative hypothesis, and above 10 indicates strong evidence.

## Results

In Table 2 an overview is given of the mean scores with standard deviations and minima and maxima for the tests used in this study.

Table 2

*Descriptive Statistics of the Main Variables of the Study*

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Range
IQ (averaged z-score)	166	0.00	0.80	-1.86 – 1.98
CITO Ability Score (0 - 150)	166	86.49	14.57	35 – 133
MCT (z-score)	162	0.00	0.91	-2.03 – 3.03
Numberline (linear fit 0.00 – 1.00)	138	.91	.10	.47 – .99
Comparison NS (% correct 0-1)	140	.63	.12	.14 – .81
Comparison NS (rt)	140	1.16	0.36	0.26 – 2.07
Comparison S (% correct 0-1)	139	.91	.08	.48 – 1.00
Comparison S (rt)	139	1.01	0.22	0.52 – 1.67
Visual spatial WM (proportion 0-1)	165	.72	.10	.36 – .94
Verbal WM (proportion 0-1)	165	.56	.08	.26 – .79
Creativity (raw score 0 - 72)	150	23.19	8.38	5 – 51

*Note.* NS = Nonsymbolic; rt = Reaction Time (in seconds); S = Symbolic; WM = Working Memory; prop = proportion score.

Table 3 shows the correlations between the numerical variables. It is noteworthy that the reaction time and accuracy of both comparison tasks were positively related, children who used more time to answer, also answered more items correctly. It is remarkable that the different number sense tasks are not strongly related. Apparently, these tasks do tap different aspects of number sense. The numberline task showed the strongest correlation with the math tasks. The symbolic comparison task also correlated with the math tasks, but only reaction time and not accuracy. This may be explained by the small variation in accuracy (91.4% of the children showed an accuracy rate of above 80%). A reverse pattern is visible for the non-symbolic task, in which accuracy was related to math, but not reaction time, although these correlations were low (BFs of 2.22 for CITO

and 1.56 for MCT). It was tested whether the symbolic tasks were more strongly related to the math task than the nonsymbolic tasks ( $\mu_{NS} < \mu_S, \mu_{NL}$ ). There was little support for this hypothesis in the data regarding the MCT (BF = 2.15, PMP = .68) and some support regarding the CITO (BF = 2.98, PMP = .75).

Table 3

*Correlations Between Variables*

Variable	1	2	3	4	5	6	7	8	9	10	11
1. IQ	—	.41***	.20*	.17	.02	.02	-.28**	.17	.19	.59***	.53**
2. Verbal WM		—	.49***	.01	.07	.02	-.10	.14	.15	.46***	.32***
3. Visual spatial WM			—	.01	.09	.13	-.09	.10	.20*	.33***	.23**
4. Comparison NS ACC				—	.30***	.13	-.11	.21*	.05	.19	.17
5. Comparison NS RT					—	.28***	.29***	.13	.08	.01	-.08
6. Comparison S ACC						—	.45***	.13	.04	.07	.02
7. Comparison S RT							—	.04	-.14	-.27**	-.21*
8. Numberline								—	.02	.39***	.24**
9. Creativity									—	.29***	.28***
10. CITO										—	.69***
11. MCT											—

*Note.* NS = Nonsymbolic; ACC = Accuracy; RT = Reaction Time; S = Symbolic; WM = Working Memory; MCT = Mathematical Creativity Test.

\*BF > 3. \*\*BF > 10. \*\*\*BF > 30.

Table 3 also shows the correlations between math and working memory. There is some evidence that the correlation between verbal working memory and academic math is higher than between visual spatial working memory and math (BF = 1.61, PMP = .62), and little in the creative math test (BF = 1.29; PMP = .56). Working memory is related to both math tasks, especially to the academic math tasks.

Creativity is also related to both math tasks, but there is no difference in strength between the two correlations. The hypothesis that creativity is more strongly linked to the creative math task than to the academic math task, was not confirmed when controlled for IQ (BF = 1.15, PMP = .13), because the hypothesis that the effects were the same received much more support (BF = 6.80, PMP = .76). Neither working memory nor creativity was related to any of the number sense tasks.

In a comparison of regression analyses (see Table 4), it was investigated whether creativity is a predictor of mathematics in addition to working memory and number sense, when corrected for IQ. Three models were compared, in the first model it was hypothesized that both working memory tasks were positively related to mathematics, but not number sense and creativity, in the second model numberline and symbolic comparison (which are most relevant, based on Table 3) were added, and in the third model creativity was added. The model including creativity was found to fit the data best. Model 3 (including creativity) received twice as much support from the data as Model 2 (without creativity), both for the academic and the creative math task. In Table 5 the estimations of the regression coefficients are given. The academic math task was – after controlling for IQ – best predicted by verbal working memory and the numberline task. For the creative math task, verbal working memory, the numberline task and the creativity task were the strongest predictors.

Table 4

*Bayesian Regression Models With IQ Included as Covariate*

Model	CITO		MCT	
	BF	PMP	BF	PMP
Unconstrained model	1.00	.01	1.00	.04
Model 1 (WM)	3.52	.04	2.44	.10
Model 2 (WM+NS)	31.83	.33	6.77	.29
Model 3 (WM+NS+creativity)	<b>58.96</b>	<b>.62</b>	<b>13.31</b>	<b>.57</b>

*Note.* Bold values indicate BFs and PMPs of models that received most support from the data. WM = Working Memory; NS = Number sense; MCT = Mathematical Creativity Test.

Table 5

*Posterior Estimations of Regression Coefficients of Model 3*

Predictor	CITO	MCT
Intelligence <sup>a</sup>	.39	.42
Visual spatial WM	.08	.05
Verbal WM	.24	.11
Numberline	.29	.15
Comparison	.10	.04
Creativity	.11	.15

*Note.* WM = Working Memory; MCT = Mathematical Creativity Test.

<sup>a</sup>Variable included as covariate in the model.

Finally, the differences between the three performance groups were tested (see Tables 6 and 7). For the domain-specific variables (the number sense task) it was expected that the MLD group performs lower than the TD group, and that the MG group would perform either higher or at the same level as the TD group. These hypotheses were compared to a model with no differences between the groups and an unconstrained model. In line with the low correlations between the nonsymbolic task and mathematics, group differences were found neither on reaction time nor on accuracy of the nonsymbolic task. For the symbolic comparison group the results are less straightforward, because Model 1 and 3 received the same amount of support regarding accuracy, and regarding reaction time the unconstrained model received most support. Since the PMP's for both accuracy and reaction time are almost as high as the most favourite model, the conclusion can be drawn that there are no relevant differences between the groups. The MLD group did score lower on the numberline task than the TD and MG groups, but the MG group did not clearly differ from the TD group.

Table 6

Means and Standard Deviations of the TD, MLD and MG Group

Variable	TD (N = 130)		MLD (N = 16)		MG (N = 20)	
	M	SD	M	SD	M	SD
IQ	0.02	0.77	-0.65	0.42	0.69	0.75
CITO	85.64	11.01	67.19	12.38	107.40	10.34
MCT	-0.11	0.78	-0.74	0.41	1.25	0.82
Numberline	0.91	0.09	0.80	0.15	0.94	0.07
Comp NS ACC	26.67	5.01	28.00	4.10	28.41	5.52
Comp NS RT	1.18	0.38	1.07	0.28	1.14	0.29
Comp S ACC	29.99	2.61	29.30	3.47	30.29	3.04
Comp S RT	1.03	0.22	0.97	0.23	0.92	0.20
Visual spatial WM	0.72	0.10	0.67*	0.09	0.75	0.08
Verbal WM	0.56	0.08	0.53	0.09	0.58	0.08
Creativity	22.94	8.49	20.14	6.51	27.75	7.54

Note. NS = Nonsymbolic; ACC = Accuracy; RT = Reaction Time; S = Symbolic; WM = Working Memory; MCT = Mathematical Creativity Test.

Table 7

Comparison of Groups on Domain-Specific Variables, Corrected for IQ

Variable	Model 0		Model 1		Model 2		Model 3	
	$\mu_{TD} > \mu_{MLD} > \mu_{MG}$		$\mu_{TD} = \mu_{MLD} = \mu_{MG}$		$\mu_{MLD} < \mu_{TD} < \mu_{MG}$		$\mu_{MLD} < \mu_{TD} = \mu_{MG}$	
	BF	PMP	BF	PMP	BF	PMP	BF	PMP
Comp NS ACC	1	.22	<b>2.39</b>	<b>.54</b>	0.49	.11	0.58	.13
Comp NS RT	1	.12	<b>3.69</b>	<b>.48</b>	1.14	.14	2.11	.26
Comp S ACC	1	.08	<b>3.99</b>	<b>.34</b>	3.07	.26	<b>3.79</b>	<b>.32</b>
Comp S RT	<b>1</b>	<b>.44</b>	<b>0.89</b>	<b>.39</b>	0.23	.10	0.16	.07
Numberline	1	.10	0.04	.00	<b>4.47</b>	<b>.44</b>	<b>4.56</b>	<b>.45</b>

Note. Bold values indicate BFs and PMPs of models that received most support from the data. NS = Nonsymbolic; ACC = Accuracy; RT = Reaction Time; S = Symbolic.

For the domain general variables, the hypothesis that the MLD group scored below the TD group and the MG group above the TD group was compared to a model in which no differences between the group exist and an unconstrained model (see Table 8). This hypothesis was confirmed for visual spatial memory. The MLD group scored lower than the TD group and the MG group scored higher on this task. However, for verbal working memory this hypothesis was not confirmed, Model 1 received most support from the data. An extra analysis showed that without a correction for IQ, Model 2 did get most support, but when correcting for IQ, no group differences were found (most support for Model 1).

Table 8

Comparison of Groups on Domain-General Variables, Corrected for IQ

Variable	Model 0		Model 1		Model 2		Model 3	
	$\mu_{TD}, \mu_{MLD}, \mu_{MG}$		$\mu_{TD} = \mu_{MLD} = \mu_{MG}$		$\mu_{MLD} < \mu_{TD} < \mu_{MG}$		$\mu_{MLD} = \mu_{TD} < \mu_{MG}$	
	BF	PMP	BF	PMP	BF	PMP	BF	PMP
IQ	1	.14	0.00	.00	<b>5.92</b>	<b>.86</b>		
MCT <sup>a</sup>	1	.14	0.00	.00	<b>6.03</b>	<b>.86</b>		
CITO <sup>a</sup>	1	.15	0.00	.00	<b>5.51</b>	<b>.85</b>		
Visual spatial WM <sup>a</sup>	1	.14	2.02	.29	<b>4.02</b>	<b>.57</b>		
Verbal WM <sup>a</sup>	1	.10	<b>7.11</b>	<b>.74</b>	1.49	.15		
Creativity <sup>a</sup>	1	.09	1.80	.16	3.77	.33	<b>4.73</b>	<b>.42</b>

Note. Bold values indicate BFs and PMPs of models that received most support from the data. MCT = Mathematical Creativity Test; WM = Working Memory.

<sup>a</sup>Controlled for IQ.

For creativity, an extra hypothesis was tested, that the MLD group does not differ from the TD group, but that the MG group outperforms both other groups. Model 3, no differences between the MLD and TD group, but higher scores in the MG group, received most support from the data.

## Discussion

In this study, the role of domain-specific (number sense) and domain-general (working memory and creativity) factors in academic and creative mathematics ability of primary school children was investigated. First it was investigated how different aspects of number sense and working memory are related to mathematical ability in fourth graders. The results confirmed our hypothesis that all three number sense tasks were related to mathematics (cf. De Smedt et al., 2009; Friso-Van den Bos, Van Luit, et al., 2015; Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010; LeFevre et al., 2010; Kolkman et al., 2013). Especially the numberline task ( $R^2$ ) and the symbolic comparison task (reaction time) were found to be good predictors of mathematical skills, even in fourth grade students. Furthermore, we found both verbal and visual spatial working memory to be predictors of math ability, with verbal working memory the strongest predictor, as expected at this age at which children start memorizing basic mathematical knowledge and procedures and because the problems in the tests were presented in a verbal format (cf. Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2015; Van der Ven et al., 2013). These results confirm earlier research that both number sense and working memory are important in learning mathematics (cf. Kroesbergen & Van Dijk, 2015; Toll et al., 2016). Of course, there are other domain general factors that influence mathematics, such as processing speed (e.g., Peterson et al., 2017; Willcutt et al., 2013), phonological skills (e.g., Barnes et al., 2014; Slot, Van Viersen, De Bree, & Kroesbergen, 2016), or other executive functions as inhibition and shifting (Friso-Van den Bos et al., 2013; Van der Ven, Kroesbergen, et al., 2012). It is not yet clear how these variables are related to creativity, and further research could take these into account to get a more comprehensive understanding of the underlying factors of mathematics.

The focus of our study, however, was on the specific role that creativity has in mathematics, because this has not been studied before in relation to other cognitive predictors of mathematics. It was investigated how creativity was related to academic math tasks and to creative math tasks. Creativity showed a medium correlation to both math tasks. Even when corrected for intelligence, creativity was not more strongly related to the creative math tasks than to the academic math task. In this study, intelligence was estimated based on two subtests of a groupwise administered intelligence test, to control for a domain general effect underlying both creativity and mathematics. Intelligence was indeed found to be the strongest predictor of mathematics. The relation between intelligence and creativity, however, was relatively low in this study. Although it was comparable to correlations found in other studies (see [Kim, 2008](#) for a meta-analysis), it was lower than in some recent studies that used broader measures of both intelligence and creativity (e.g. [Benedek, Franz, Heene, & Neubauer, 2012](#); [Silvia, 2008](#)). More research is needed to investigate the possible dependency between creativity and intelligence, and how this may influence mathematics. A closer look at different aspects of intelligence and the use of individually administered tests are recommended.

Furthermore, it was found that a regression analysis including creativity better fits the data than a regression with only IQ, working memory and number sense included. These results confirm our hypothesis that creativity is an important domain-general factor in mathematics that deserves more attention in education ([Kattou et al., 2013](#); [Schoevers et al., 2017](#)). In future research, this relation should receive more attention. Although we did not find differences in relations between the creative and the academic math test, we do think further research is necessary. The academic math test seems to rely more on memorized knowledge and procedures and less on creative problem solving than the creative math test, but still significant relations with creativity were found. For future research it would therefore be interesting to include pure measures of memorized math knowledge, to test whether creativity is indeed more weakly related (or not at all) to such a test. Accordingly, other types of creative math tests could also be used in future research. The creative math test used in this study required creative thinking based on knowledge of arithmetical operations and number relations which were possessed by most students ([Schoevers et al., 2017](#)). More difficult creative math questions might lead to larger role of creativity in creative math tasks than in academic math tasks.

Another aspect that should be taken into account is that we have used one single measure of creativity. Since creativity is a complex, multidimensional construct ([Cropley, 2010](#)), it is advised to take a broader set of measures into account. The test that is used in this study, is based on a holistic view on creativity. Still, the parallels between task requirements in the creativity test and problem solving are evident: Children who use many elements or add new elements, who make many connections between elements and make unconventional combinations, score highest on the creativity test. In mathematical problem solving, the discovery of solutions takes also place by the activation of multiple sources of information and new ideas, and by making new and unconventional combination between those ideas ([Hadamard, 1996](#); [Koichu et al., 2006](#)). So, although the creativity and the mathematical tasks differ in task format regarding output (drawing versus mathematical reasoning), the same underlying construct of creativity seems to play a role. However, it should be noted that the results of this study only show that creativity and math are related, but interpretations of the underlying cognitive processes are not possible. For example, some have argued that the differences between intelligence and creativity may be smaller than often thought, and that creativity relies so much on executive processes, that it is questionable whether it should be regarded a separate ability ([Silvia, 2015](#)). Future studies with experimental designs and the use of online measures such as eye tracking or EEG could give more insight in the exact role of creativity in learning mathematics.

Thirdly, we have investigated differences between MLD, MG, and TD children on number sense, working memory and creativity. Our hypothesis that children with MLD have deficits in visual spatial working memory and number sense was confirmed (cf. Ashkenazi et al., 2013; Geary et al., 2007; Kroesbergen & Van Dijk, 2015). Although, in general, verbal working memory was found to be as important as visual working memory in fourth graders, the difference between MLD and TD children is especially visible in visual working memory, which is in accordance with former literature (Ashkenazi et al., 2013; Szucs et al., 2013). This finding may be explained by the same mechanism that explains why visual working memory is more important in elementary math, while verbal working memory is more involved in later math (Van der Ven et al., 2013; Van de Weijer-Bergsma, Kroesbergen, & Van Luit, 2015): Children with MLD have difficulties with automatizing facts and procedures (e.g., Andersson, 2008; Geary et al., 2012), which leads to math problems being relatively 'new' for them, because they can't rely on memorized facts and procedures. Visual spatial working memory seems to be more involved in new problems, and thus in the problem solving of children with MLD. Another explanation is that verbal working memory is more related to IQ, and by controlling for IQ, the differences between the groups disappeared.

Furthermore, children with MLD were not found to differ from TD children in creativity. An interesting finding from this study is that MG children did not significantly differ from TD children in number sense, but that they did differ in visual spatial working memory and creativity. It seems that a certain level of number sense is needed to perform well in mathematics, but that above this threshold other factors, such as creativity, come into play. It should be noted of course that the sample size in this study was relatively small, which make generalizations difficult. It is advised to further investigate the differences between MG and TD children with larger samples. Furthermore, both creativity and mathematical ability may be influenced by group factors. When this study is replicated, larger samples would also make multilevel analyses possible. Another factor that should be taken into account is that this study does not tell us how we should interpret the directions of the relation between creativity and mathematics. Based on literature, we assume that creativity supports mathematical thinking and problem solving during mathematical tasks, but the experience with mathematical problem solving also might influence children's creative skills, and other factors may also have a mediating or moderating role in these relations. This is one of the first studies that examined relations between creativity and mathematics in relation to working memory and number sense, but further research is necessary to give insight in the direction of these relations.

To conclude, symbolic number sense, working memory and creativity are all related to academic and creative mathematical ability. Furthermore, groups with math learning disabilities (MLD) and mathematical giftedness (MG) were compared to a typically developing (TD) group. The results show that the MLD group scores significantly lower on number line estimation and visual spatial working memory than the TD group, while the differences between the MG group and the TD group were found in visual spatial working memory and creativity. It is concluded that visual spatial working memory is a domain general factor that discriminates between the three ability groups, that number sense discriminates between MLD and TD children and that creativity discriminates between TD and MG children. This study confirms the hypothesis that creativity plays a significant role in mathematics.

## Funding

The authors have no funding to report.

## Competing Interests

The authors have declared that no competing interests exist.

## Acknowledgments

The authors have no support to report.

## References

- Allison, P. D. (2009). Missing data. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology* (pp. 72-89). London, United Kingdom: Sage.
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the automated working memory assessment. *Educational Psychology, 28*, 725-734. doi:10.1080/01443410802243828
- Andersson, U. (2008). Mathematical competencies in children with different types of learning difficulties. *Journal of Educational Psychology, 100*, 48-66. doi:10.1037/0022-0663.100.1.48
- Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W., Swigart, A. G., & Menon, V. (2013). Visuo-spatial working memory is an important source of domain-general vulnerability in the development of arithmetic cognition. *Neuropsychologia, 51*, 2305-2317. doi:10.1016/j.neuropsychologia.2013.06.031
- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829-839. doi:10.1038/nrn1201
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York, NY, USA: Academic Press.
- Barnes, M. A., Raghobar, K. P., English, L., Williams, J. M., Taylor, H., & Landry, S. (2014). Longitudinal mediators of achievement in mathematics and reading in typical and atypical development. *Journal of Experimental Child Psychology, 119*, 1-16. doi:10.1016/j.jecp.2013.09.006
- Benedek, M., Franz, F., Heene, M., & Neubauer, A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences, 53*, 480-485. doi:10.1016/j.paid.2012.04.014
- Bergman-Nutley, S., & Klingberg, T. (2014). Effect of working memory training on working memory, arithmetic and following instructions. *Psychological Research, 78*, 869-877. doi:10.1007/s00426-014-0614-0
- Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics, 58*, 45-75. doi:10.1007/s10649-005-0808-x
- Cropley, A. J. (2010). Creativity in the classroom: The dark side. In D. H. Cropley (Ed.), *The dark side of creativity* (pp. 297-315). doi:10.1017/CBO9780511761225.016

- Dark, V. J., & Benbow, C. P. (1991). Differential enhancement of working memory with mathematical versus verbal precocity. *Journal of Educational Psychology, 83*, 48-60. doi:10.1037/0022-0663.83.1.48
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42. doi:10.1016/0010-0277(92)90049-N
- 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*, 469-479. doi:10.1016/j.jecp.2009.01.010
- Friso-Van den Bos, I., Kroesbergen, E. H., & Van Luit, J. E. H. (2014). Number sense in kindergarten children: Factor structure and working memory predictors. *Learning and Individual Differences, 33*, 23-29. doi:10.1016/j.lindif.2014.05.003
- Friso-Van den Bos, I., Schoevers, E. M., Slot, E. M., & Kroesbergen, E. H. (2015). *The Dutch Assessment of Number Sense (DANS): Analyses of the conditions of the number line, symbolic comparison, and non-symbolic comparison task*. Department of Education and Pedagogy, Utrecht University, Utrecht, The Netherlands.
- 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
- Friso-Van den Bos, I., Van Luit, J. E. H., Kroesbergen, E. H., Xenidou-Dervou, I., van Lieshout, E. C. D. M., Van der Schoot, M., & Jonkman, L. M. (2015). Pathways of number line development in children: Predictors and risk for adverse mathematical outcome. *Zeitschrift für Psychologie mit Zeitschrift für Angewandte Psychologie, 223*, 120-128. doi:10.1027/2151-2604/a000210
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*, 1520-1533. doi:10.1111/j.1467-8624.2010.01489.x
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., . . . Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology, 46*, 1731-1746. doi:10.1037/a0020662
- Geake, J. G. (2009). Neuropsychological characteristics of academic and creative giftedness. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 261-273). Berlin, Germany: Springer.
- 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., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The number sets test. *Journal of Psychoeducational Assessment, 27*, 265-279. doi:10.1177/0734282908330592
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78*, 1343-1359. doi:10.1111/j.1467-8624.2007.01069.x

- Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology, 104*, 206-223. doi:10.1037/a0025398
- Gill, J. (2008). *Bayesian methods: A social and behavioral sciences approach* (2nd ed.). Boca Raton, FL, USA: Chapman and Hall/CRC.
- Hadamard, J. (1996). *The mathematician's mind: The psychology of invention in the mathematical field*. Princeton, NJ, USA: Princeton University Press.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*, 23-34. Retrieved from <http://www.tqmp.org/RegularArticles/vol08-1/p023/p023.pdf>doi:10.20982/tqmp.08.1.p023
- HCO. (2016). *Tabellen tussenopbrengsten CITO LVS* [Tables of CITO LVS]. Retrieved from [http://www.hco.nl/images/paginas/werkenaanopbrengsten/OGW\\_en\\_data-analyse/2016\\_Tabellen\\_tussenopbrengsten\\_groepsniveau\\_CITO\\_LOVS.pdf](http://www.hco.nl/images/paginas/werkenaanopbrengsten/OGW_en_data-analyse/2016_Tabellen_tussenopbrengsten_groepsniveau_CITO_LOVS.pdf)
- Hojtink, H. E. (2012). *Informative hypotheses: Theory and practice for behavioral and social scientists*. Boca Raton, FL, USA: CRC Press.
- Hojtink, H. E., Klugkist, I. E., & Boelen, P. A. (2008). *Bayesian evaluation of informative hypotheses: Statistics for social and behavioral sciences*. New York, NY, USA: Springer Science+Business Media.
- Holmes, J., & Gathercole, S. E. (2014). Taking working memory training from the laboratory into schools. *Educational Psychology, 34*, 440-450. doi:10.1080/01443410.2013.797338
- Janssen, J., Scheltens, F., & Kraemer, J. (2007). *Rekenen-wiskunde: Handleiding* [Mathematics: Manual]. Arnhem, The Netherlands: Cito.
- Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke verantwoording van de toetsen LOVS rekenenwiskunde voor groep 3 tot en met 8* [Scientific justification of the mathematics test for Grade 1 to Grade 6]. Arnhem, The Netherlands: Cito.
- JASP Team. (2016). *JASP* (Version 0.8.0.0) [Computer software]. Retrieved from <https://jasp-stats.org>
- 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*, 82-88. doi:10.1016/j.lindif.2009.07.004
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773-795. doi:10.1080/01621459.1995.10476572
- Kattou, M., Kontoyianni, K., Pitta-Pantazi, D., & Christou, C. (2013). Connecting mathematical creativity to mathematical ability. *ZDM, 45*, 167-181. doi:10.1007/s11858-012-0467-1
- Kaufman, J. C., & Sternberg, R. J. (2010). *Cambridge handbook of creativity*. Cambridge, United Kingdom: Cambridge University Press.
- Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *The Journal of Creative Behavior, 42*, 106-130. doi:10.1002/j.2162-6057.2008.tb01290.x

- Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London, United Kingdom: Routledge.
- Koichu, B., Berman, A., & Moore, M. (2006). Patterns of middle school students' heuristic behaviors in solving seemingly familiar problems. In J. Novotna, H. Moraova, M. Kratka, & N. Stehlikova (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 457-464). Prague, Czech Republic: Charles University.
- Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25*, 95-103. doi:10.1016/j.learninstruc.2012.12.001
- Kluytmans, A., Van De Schoot, R., Mulder, J., & Hoijtink, H. (2012). Illustrating Bayesian evaluation of informative hypotheses for regression models. *Frontiers in Psychology, 3*, Article 2. doi:10.3389/fpsyg.2012.00002
- Kroesbergen, E. H., & Van Dijk, M. (2015). Working memory and number sense as predictors of mathematical (dis-)ability. *Zeitschrift für Psychologie mit Zeitschrift für Angewandte Psychologie, 223*, 102-109. doi:10.1027/2151-2604/a000208
- Kroesbergen, E. H., Van 't Noordende, J. E., & Kolkman, M. E. (2014). Training working memory in kindergarten children: Effects on working memory and early numeracy. *Child Neuropsychology, 20*, 23-37. doi:10.1080/09297049.2012.736483
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174. doi:10.2307/2529310
- LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development, 81*, 1753-1767. doi:10.1111/j.1467-8624.2010.01508.x
- Leikin, R. (2009). Exploring mathematical creativity using multiple solution tasks. In R. Leikin, A. Berman, & B. Koichu (Eds.), *Creativity in mathematics and the education of gifted students* (pp. 129-145). Rotterdam, The Netherlands: Sense Publishers.
- Leikin, R., Koichu, B., & Berman, A. (2009). Mathematical giftedness as a quality of problem-solving acts. In R. Leikin, A. Berman, & B. Koichu (Eds.), *Creativity in mathematics and the education of gifted students* (pp. 115-127). Rotterdam, The Netherlands: Sense.
- Leikin, R., & Lev, M. (2007, July). Multiple solution tasks as a magnifying glass for observation of mathematical creativity. In J. H. Woo, H. C. Lew, K. S. Park, & D. Y. Seo (Eds.), *Proceedings of the 31st international conference for the psychology of mathematics education* (Vol. 3, pp. 161-168). Seoul, Korea: The Korea Society of Educational Studies in Mathematics.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198-1202. doi:10.1080/01621459.1988.10478722
- Mann, E. L. (2006). Creativity: The essence of mathematics. *Journal for the Education of the Gifted, 30*, 236-260. doi:10.4219/jeg-2006-264
- Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*, 1224-1237. doi:10.1111/j.1467-8624.2011.01608.x

- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270-291. doi:10.1037/a0028228
- Mulder, J., Hoijtink, H., & De Leeuw, C. (2012). BIEMS: A Fortran 90 program for calculating Bayes factors for inequality and equality constrained models. *Journal of Statistical Software, 46*(2), 1-39. doi:10.18637/jss.v046.i02
- Mussolin, C., Mejias, S., & Noël, M. P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition, 115*, 10-25. doi:10.1016/j.cognition.2009.10.006
- Passolunghi, M. C., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *The British Journal of Educational Psychology, 82*, 42-63. doi:10.1111/j.2044-8279.2011.02039.x
- Peterson, R. L., Boada, R., McGrath, L. M., Willcutt, E. G., Olson, R. K., & Pennington, B. F. (2017). Cognitive prediction of reading, math, and attention: Shared and unique influences. *Journal of Learning Disabilities, 50*, 408-421. doi:10.1177/0022219415618500
- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., . . . Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*, 33-41. doi:10.1016/j.cognition.2010.03.012
- Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*, 83-96. doi:10.1207/s15326985ep3902\_1
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110-122. doi:10.1016/j.lindif.2009.10.005
- Sasanguie, D., Göbel, 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*, 418-431. doi:10.1016/j.jecp.2012.10.012
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA, USA: Psychology Software Tools Inc.
- Schoevers, E. M., Kroesbergen, E. H., & Kattou, M. (2017). *Primary school mathematics requires more than facts and procedures: The role of creativity in fourth-grade math learning*. Manuscript submitted for publication.
- Schuchardt, K., Maehler, C., & Hasselhorn, M. (2008). Working memory deficits in children with specific learning disorders. *Journal of Learning Disabilities, 41*, 514-523. doi:10.1177/0022219408317856
- Silvia, P. J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences, 44*, 1012-1021. doi:10.1016/j.paid.2007.10.027
- Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychology Review, 27*, 599-606. doi:10.1007/s10648-015-9299-1

- Slot, E. M., Van Viersen, S., De Bree, E. H., & Kroesbergen, E. H. (2016). Shared and unique risk factors underlying mathematical disability and reading and spelling disability. *Frontiers in Psychology, 7*, Article 803. doi:10.3389/fpsyg.2016.00803
- Söderqvist, S., & Bergman-Nutley, S. (2015). Working memory training is associated with long term attainments in math and reading. *Frontiers in Psychology, 6*, Article 1711. doi:10.3389/fpsyg.2015.01711
- Sriraman, B. (2005). Are giftedness and creativity synonyms in mathematics? *Journal of Secondary Gifted Education, 17*, 20-36. doi:10.4219/jsge-2005-389
- Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex, 49*, 2674-2688. doi:10.1016/j.cortex.2013.06.007
- Toll, S. W. M., Kroesbergen, E. H., & Van Luit, J. E. H. (2016). Visual working memory and number sense: Testing the double deficit hypothesis in mathematics. *The British Journal of Educational Psychology, 86*, 429-445. doi:10.1111/bjep.12116
- Urban, K. K. (2005). Assessing creativity: The Test for Creative Thinking - Drawing Production (TCT-DP). *International Education Journal, 6*, 272-280. Retrieved from <http://files.eric.ed.gov/fulltext/EJ854980.pdf>
- Urban, K. K., & Jellen, H. G. (1996). *Test for Creative Thinking - Drawing Production (TCT-DP)*. Lisse, The Netherlands: Swets & Zeitlinger.
- Van der Ven, S. H. G., Boom, J., Kroesbergen, E. H., & Leseman, P. P. M. (2012). Microgenetic patterns of children's multiplication learning: Confirming the Overlapping Waves model by latent growth modeling. *Journal of Experimental Child Psychology, 113*, 1-19. doi:10.1016/j.jecp.2012.02.001
- Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematical skills: A dynamic relationship. *The British Journal of Educational Psychology, 82*, 100-119. doi:10.1111/j.2044-8279.2011.02035.x
- Van der Ven, S. H. G., Van der Maas, H. L. J., Straatemeier, M., & Jansen, B. R. J. (2013). Visuospatial working memory and mathematical ability at different ages throughout primary school. *Learning and Individual Differences, 27*, 182-192. doi:10.1016/j.lindif.2013.09.003
- Van de Schoot, R., Hoijting, H., Mulder, J., Van Aken, M. A. G., Orobio de Castro, B., Meeus, W., & Romeijn, J. (2011). Evaluating expectations about negative emotional states of aggressive boys using Bayesian model selection. *Developmental Psychology, 47*, 203-212. doi:10.1037/a0020957
- Van de Weijer-Bergsma, E., Kroesbergen, E. H., Jolani, S., & Van Luit, J. E. H. (2016). The Monkey game: A computerized verbal working memory task for self-reliant administration in primary school children. *Behavior Research Methods, 48*, 756-771. doi:10.3758/s13428-015-0607-y
- Van de Weijer-Bergsma, E., Kroesbergen, E. H., Prast, E. J., & Van Luit, J. E. H. (2015). Validity and reliability of an online visual spatial working memory task for self-reliant administration in school-aged children. *Behavior Research Methods, 47*, 708-719. doi:10.3758/s13428-014-0469-8

- Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Verbal and visual-spatial working memory and mathematical ability in different domains throughout primary school. *Memory & Cognition*, *43*, 367-378. doi:[10.3758/s13421-014-0480-4](https://doi.org/10.3758/s13421-014-0480-4)
- Van Dijk, H., & Tellegen, P. J. (2004). *Handleiding NIO* [Manual of the NIO]. Amsterdam, The Netherlands: Boom test uitgevers.
- Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., & Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities*, *46*, 500-516. doi:[10.1177/0022219413477476](https://doi.org/10.1177/0022219413477476)