MENTAL ROTATION AND MENTAL FOLDING IN 7- AND 8-YEAR-OLD CHILDREN

Laura Ilen

Master's thesis

Department of Psychology

University of Jyväskylä

December 2016

UNIVERSITY OF JYVÄSKYLÄ

Department of Psychology

ILEN, LAURA: Mental rotation and mental folding in 7- and 8-year-old children

Master's thesis, 39 pp.

Supervisors: Pekka Räsänen and Jarkko Hautala

Psychology

December 2016

Mental rotation and mental folding are mental transformations of spatial representations. Mental rotation changes the orientation of an object, whereas mental folding requires picturing how an object will look like after it has been folded. The skills have been studied a lot but still it is not sure if they are similar or different processes. Earlier research has found a connection between mental rotation and mathematical skills but a relation between mathematics and mental folding has not been proved. However, most of the research has been done with adult participants, so the mental transformations in children have not been studied as much. The goal of the present study was to examine the mental transformation processes in 7- and 8-year-old children, and to test whether the processes can predict the performance in mathematics. The participants were 118 children from the first and second grades of the University of Jyväskylä Teacher Training School. We used new computer-based tests to test the mental rotation and mental folding skills, and paper-and-pencil tests to measure reading, mathematics and visuospatial skills. The results showed that the mental rotation and mental folding tasks correlated with each other, but they cannot be completely similar processes since the correlations with paper-and-pencil tests differed. In addition, only mental rotation had a connection with mathematical skills, although it predicted math scores weakly. Furthermore, only mental folding showed gender differences: girls had better scores in mental folding task than boys. However, the results showed that most of the children at this age did not mentally rotate images. As some of the task reliabilities were quite low, some improvements to the next test version are needed to get more reliable results of the mental transformation processes at this age.

KEYWORDS: mental rotation, mental folding, development, mathematical skills, sex differences

JYVÄSKYLÄN YLIOPISTO

Psykologian laitos

ILEN, LAURA: Mielessä kääntäminen ja mielessä taittelu 7- ja 8-vuotiailla lapsilla

Pro gradu –tutkielma, 39 s.

Ohjaajat: Pekka Räsänen ja Jarkko Hautala

Psykologia

Joulukuu 2016

Mielessä kääntäminen ja mielessä taittelu ovat avaruudellisten representaatioiden mielessä muuntamista. Mielessä kääntäminen muuttaa objektin suunnan, kun taas mielessä taittelussa täytyy kuvitella, miltä objekti näyttää taittamisen jälkeen. Taitoja on tutkittu paljon, mutta silti ei ole varmuutta, ovatko ne samanlaisia vai erilaisia prosesseja. Aiemmat tutkimukset ovat löytäneet yhteyden mielessä kääntämisen ja matematiikan taitojen välille, mutta mielessä taittelun yhteyttä matematiikkaan ei ole osoitettu. Kuitenkin, suurin osa tutkimuksesta on tehty aikuisilla, joten lasten mielessä muuntamisen taitoja ei ole tutkittu yhtä laajasti. Tässä tutkimuksessa tavoitteena oli tutkia mielessä kääntämisen ja mielessä taittelun prosesseja 7- ja 8-vuotiailla lapsilla, ja selvittää, voivatko prosessit ennustaa matematiikassa suoriutumista. Tutkimukseen osallistui 118 lasta Jyväskylän Normaalikoulun ensimmäiseltä ja toiselta luokalta. Tutkimuksessa käytettiin uutta tietokonetestistöä testaamaan lasten mielessä kääntämisen ja mielessä taittelun taitoja sekä kynä-paperitestejä selvittämään lukemista, matematiikkaa ja visuospatiaalisia taitoja. Tulokset osoittivat, että mielessä kääntämisen ja mielessä taittelun tehtävät korreloivat keskenään, mutta prosessit eivät voi olla täysin samanlaisia, sillä korrelaatiot kynä-paperitesteihin erosivat. Vain mielessä kääntäminen oli yhteydessä matematiikkaan, mutta se ennusti matematiikan tuloksia heikosti. Lisäksi, sukupuolieroja esiintyi vain mielessä taittelun tehtävässä, jossa tytöt suoriutuivat poikia paremmin. Kuitenkin, tulokset osoittavat, että suurin osa tämän ikäisistä lapsista ei kääntänyt kuvia mielessään. Koska joidenkin tehtävien reliabiliteetit olivat melko matalia, joitakin muutoksia seuraavaan testiversioon on tehtävä, jotta saataisiin luotettavampia tuloksia mielessä muuntamisen prosesseista tässä iässä.

AVAINSANAT: mielessä kääntäminen, mielessä taittelu, kehitys, matematiikan taidot, sukupuolierot

TABLE OF CONTENTS

1. INTRODUCTION

Mental rotation and mental folding are both mental transformations of spatial representations. Mental transformations involve imagining and mentally transforming objects, and they are commonly measured abilities when researching and defining spatial skills. Understanding mental transformations is thus important to the wider understanding, assessment and rehabilitation of spatial ability. Even though lots of research has focused on mental rotation and mental folding, it is not clear whether they are similar or different processes. Early factor analytic studies divided mental rotation and mental folding into two different factors: spatial relations and spatial visualization (Guilford, Fruchter, & Zimmerman, 1952; Lohman, 1979). Linn & Petersen's (1985) categorization of spatial abilities, commonly used by researchers, also separates mental rotation and mental folding. In their meta-analysis, mental rotation forms a separate category whereas mental folding falls into the spatial visualization category. However, it is possible that the distinction between mental rotation and mental folding has resulted from the differences in the task difficulty rather than the differences in their processes. In the latest classification of spatial abilities, both mental rotation and mental folding are seen as similar, intrinsic and dynamic processes (Uttal et al., 2013). In other words, they both require thinking about one single and dynamic object at the time. Harris, Hirsh-Pasek and Newcombe (2013) recently published a review of spatial transformations to clear up the differences and similarities between mental rotation and mental folding, and found a lot of contradicting results. The current research examines the processes of mental rotation and mental folding in children. The goal is to find out more about the processes in 7- and 8-year-old children, and examine whether we can find a relation to mathematical skills.

Mental rotation is a mental transformation of spatial representation that changes its orientation. A typical mental rotation task requires distinguishing whether two objects are the same or mirror images after the other has been rotated. Researchers are usually interested in the reaction time and the accuracy of judgments. A noted characteristic of mental rotation is the linearity of reaction time functions. Shepard and Metzler (1971) found out that the response times of mental rotation tasks increased in a strikingly linear fashion depending on the angular difference between two block figures in three-dimensional space, which has been taken as evidence that subjects solve the problem by mentally rotating images (Cooper & Shepard, 1973). The results have been replicated several times with both 3D and 2D objects and with random, angular shapes (Cooper, 1975), alphanumeric characters (Cooper & Shepard, 1973), line drawings of nameable objects (Jolicoeur, 1988), and matrix patterns (Bethell-Fox & Shepard, 1988). Mental rotation has also been studied by using line drawings of hands and other body parts as objects, and the results have showed that when subjects mentally rotate pictures of body parts, they imagine transforming their own corresponding body parts (Cooper & Shepard, 1975; Parsons, 1987a). The rotation times for hands are not strictly linear but seem to depend more of the position of stimuli in such a way that the rotation is slower if the presented picture of hand is in awkward position (Parsons, 1987b). Neuroimaging studies have shown that mental rotation of hands activates the motor areas involved in preparation of movements, whereas mental rotation of cubes figures activates parietal regions but not frontal motor regions (Kosslyn, Digirolamo, Thompson, & Alpert, 1998). Therefore, it would seem that when people imagine rotating their hands or other body parts, there is a motor element involved.

Mental folding requires picturing how an object will look like after it has been folded. Objects can be two-dimensional, like a drawing of a folded paper sheet, or three-dimensional, like a cube that has been unfolded onto a flat surface. There are many versions of mental folding tasks, and often used tests are based on Thurstone's punched holes (1938) and Shepard and Feng's unfolded cubes (1972). In Shepard and Feng's task, participants saw an unfolded cube and they had to imagine whether two arrays would meet if the squares were folded back up into the cube. The response times increased linearly (from 2 to about 15 seconds) with the sum of the number of squares involved in the task. The results are similar to the results found in mental rotation tasks. Shepard and Feng thus concluded that the majority of subjects solved tasks by mentally folding up the squares. However, the linearity of reaction times in mental folding tasks has been found only with objects of three-dimensional shape and only with adults. Neuroimaging studies show that mental folding activates the parietal cortex, like mental rotation, but the processes may occur in different area. Mental rotation may activate more right parietal cortex whereas mental folding seems to take place on bilateral parietal regions (Milivojevic, Johnson, Hamm, & Corballis, 2003).

A recent review of Harris et al. (2013) suggests that the cognitive processes behind mental rotation and mental folding are at least partly different. One difference is the rigidness; mental rotation is a rigid transformation where the object stays geometrically similar after rotation, whereas mental folding is a non-rigid transformation where the object changes after folding. Moreover, there has been a great deal of discussion about strategies behind mental transformations. Even though there is evidence that mental rotation tasks are solved spatially by forming a mental image of the figure and mentally rotating the image, other strategies can be sometimes used, too. Cooper and Shepard (1973) found shorter response times for normal than for mirrored stimuli and concluded that participants did not have to rotate the object mentally if it was near the upright. They also noticed that reaction times were faster with more familiar objects like letters and numbers. Hamm, Johnson and Corballis (2004) suggested that mirrored stimuli are rotated slower because they are flipped after plane rotation. There is evidence that the rotation of unfamiliar and complex stimuli can be eased by rotating them piece by piece (Bethell-Fox & Shepard, 1988). In addition, studies have found individual differences in use of strategy (e.g. Searle & Hamm, 2012). Khooshabeh, Hegarty and Shipley (2013) suggested that poor imagers tended to use a piecemeal strategy to rotate, whereas better imagers rotated piece by piece only when tasks were demanding. Moreover, Geiser, Lehmann and Eid (2006) even found participants who did not mentally rotate images at all. The use of analytical strategy seems to be less efficient solving mental rotation tasks (Linn & Petersen, 1985). Analytic strategy requires more reasoning than mental manipulation. Some researchers suggest that mental folding and other spatial visualization tasks can be solved using analytical instead of spatial strategy or using both strategies at the same time (Linn & Petersen, 1985; Lohman, 1979). It is yet unclear if mental folding tasks and mental rotation tasks are really solved using different strategy, or if the differences result from the difficulty of mental folding tests. It seems that using several strategies can be helpful when subjects solve complex mental folding tasks (Kyllonen, Lohman, & Snow, 1984).

1.1. Development and sex differences

Although mental rotation and mental folding have been studied a lot, the research has focused mainly on these skills in adults. Mental transformations in children have not been examined as much, so the development is not yet clear. Especially the early development of mental folding has been almost totally ignored in research. There is only one recent published study of the topic. Harris, Newcombe and Hirsh-Pasek (2013) researched 4-7-year-old children and found out that mental folding skills appear around 5.5 years of age. The results showed large individual differences though. The results are fairly similar that have been found in mental rotation. Piaget and Inhelder (1971) claimed that children are starting to be able to do mental transformations at the age of 7 or 8, when they attain the concrete operational stage. Marmor (1975; 1977) found differing results. In her studies 4- and 5-year-olds were already capable of mentally rotating images, which could be concluded from the linear trends of reaction time functions. However, Dean and Harvey (1979) challenged Marmor’s results and claimed that before the age of 7 children did not show linear trends but performed near chance, a view in agreement with Piaget and Inhelder. A recent study of Frick, Ferrara and Newcombe (2013) examined mental rotation in 3.5-5.5-year-old children and found out that accuracy increased with age. In their study, 4-year-olds performed at chance even after given experience, whereas 5-year-olds’ reaction times were more linear and they profited from the experience. Nonetheless, the individual differences and developmental progression were still notable at age 5. Also Kail, Pellegrino and Carter (1980) proved that mental rotation accelerated with development. In their study, the response time of third- and fourth-graders (7 msec/degree) was significantly slower than that of college students (less than 4 msec/degree). Some researchers have suggested that young children use more motor representations and benefit more from analogous hand movements than older children and adults when mentally rotating objects (Frick, Daum, Wilson, & Wilkening, 2009; Funk, Brugger, & Wilkening, 2005). In the study of Frick, Daum, Walser and Mast (2009) participants had to turn a wheel in one direction and mentally rotate an object to the opposite direction. Conflicting hand movements disturbed mental rotation when participants were 5- or 8-year-old children but not when they were 11-year-old children or adults. The results demonstrate that age increases the ability to distinguish motor processes from mental imagery.

Though it is known that mental rotation develops through childhood, it has been suggested that some kind of mental rotation would be possible even in infancy at the age of 4 months (Rochat & Hespos, 1996). There is also evidence that 3- to 5-month-old male infants recognize rotated normal and mirror letters better than female infants (Moore & Johnson, 2008; Quinn & Liben, 2008). The results match with a great deal of other research findings of sex differences in mental rotation. It seems that males perform better than females already at age 4 (Levine, Huttenlocher, Taylor, & Langrock, 1999) and through adulthood (Geiser et al., 2006). In addition, sex differences have been noticed to increase with age (Voyer, Voyer, & Bryden, 1995). Two meta-analyses have clarified that mental rotation shows strong sex differences in favor of males from the age of 10 onwards whereas mental folding does not show any at all (Linn & Petersen, 1985; Voyer et al., 1995). Thus, it would seem that the two tasks differ in some way and cannot be completely similar processes. Harris et al. (2013) suggested that one explanation behind sex differences might be the rigidness since there is evidence that males tend to perform better in rigid tasks, i.e. rotation, translation and mirroring. In addition, some researchers have argued that females may use more analytical strategies in solving mental rotation tasks, which can explain the lower performance (Geiser et al., 2006; Kail, Carter, & Pellegrino, 1979). However, it is still unclear why sex differences have been found in mental rotation but not in mental folding.

1.2. Relation to STEM and malleability

Spatial ability can predict success in science, technology, engineering and mathematics (STEM) (Shea, Lubinski, & Benbow, 2001). Both mental folding and mental rotation relate to entry and performance in STEM disciplines (Wai, Lubinski, & Benbow, 2009), but it is not clear if these two tasks predict success differently. Usually researchers have used a composite score, which makes it hard to find out their separate relations to STEM success. There is some evidence of the relation between mental rotation and success in mathematics, measured with the number of completed math classes (Cherney & Collaer, 2005) and math test scores (Casey, Nuttall, Pezaris, & Benbow, 1995). Similar results have not been found with mental folding, which can be a result from the lack of research, however. We do not know if mental rotation and mental folding are differently related to STEM, or if the few differences found so far have been only a consequence of differences in task complexity.

Since spatial skills are related to STEM success, it is possible that performance in these disciplines can be improved by training spatial ability (Baenninger & Newcombe, 1989). In the meta-analysis of Uttal et al. (2013) training of spatial skills was noticed to be effective, transferable and durable. The meta-analysis proved that spatial skills are malleable through the life span for both sexes. Mental rotation can be improved from childhood to adulthood with practice (Kail, 1986) and playing video games (Feng, Spence, & Pratt, 2007). Terlecki, Newcombe and Little (2008) found out that practicing video games improved mental rotation, the effect transferred to other spatial tasks, and the training and transfer effects endured for several months. Mental rotation training can also improve mental folding skills (Lizarraga & Ganuza, 2003) and vice versa (Wright, Thompson, Ganis, Newcombe, & Kosslyn, 2008). In the study of Wright et al. (2008) participants practised either mental rotation or mental folding, and in spite of practised task, both skills improved. These results implicate that there is some shared mechanism behind different mental transformations. In conclusion, mental rotation and mental folding are malleable at all ages and relate to STEM success, so the early practice can be valuable in improving children's STEM learning and have a great impact later in life as well.

1.3. Aims of the study

The current research focuses on the mental transformations of children in Grades 1 and 2. The goal is to examine the processes at this age and to find out whether we can find the relation between mathematical skills and one or both of these tasks. Since this is the first study where we use the current computer-based mental rotation and mental folding tests, one goal is to examine the reliability and validity of the tests. We want to observe the correlations among mental folding and mental rotation, and their correlations with reading, math and visuospatial skills. Mental rotation and mental folding tests should have a relation with tasks of spatial visualization and perception that require similar mental manipulation processes.

There are contradictory findings of the age in which children are capable of mentally rotate items, but it has been claimed that they would start to be able to do so at the age of 7 or 8 (Piaget & Inhelder, 1971). According to Cooper and Shepard (1973), reaction times in mental rotation tasks increase in a linear fashion depending on the angle, which proves that participants rotate items mentally. Therefore, we want to examine if we can find a linear trend in reaction times already at this age. We also know that mental rotation accelerates with development (Kail et al., 1980), so one goal is to examine whether we can find different results among first-graders and second-graders. There is not much research of mental folding in children, so the current study aims to better understanding of the processes in childhood. If differences in the results among first-graders and second-graders are found, it suggests that the processes of mental folding benefit from learning.

One goal of the current study is to examine how mental folding and mental rotation relate to mathematical thinking skills. We measure mathematical thinking dividing it into three cognitive domains depending on the complexity – knowing, applying and reasoning (Mullis & Martin, 2013). Knowing includes the concepts or facts, whereas applying requires participants to apply knowledge to solve the problems. Reasoning is the most complex process and it covers unfamiliar situations and multiphase problems. There is some evidence of the correlation among mental rotation and math scores (Casey et al., 1995) but the studies are not conducted with young children. Therefore, we aim to find out whether mental rotation correlates with these three domains of mathematical thinking at the age of 7 and 8. We are interested in finding out whether mental rotation can predict mathematical skills, and we assume that we can find a relation between mental rotation and the reasoning task in particular. In addition, we know that mental rotation of hands is a more complex process than rotation of objects since it involves a motor element (Cooper & Shepard, 1975;

Kosslyn et al., 1998), so we assume that also the mental rotation task where stimuli are drawings of hands, correlates with a complex mathematical task that requires reasoning.

We are also interested in examining the relation among mental folding and mathematical thinking, even though previous studies do not show any connection among the skills. However, the lack of findings can result from a small number of studies or the task difficulty since mental folding has often been measured with complex, multi-phased tasks. In the present study, we measured both mental rotation and mental folding skills with simple computer-based tests to eliminate possible effects of complexity. Both tasks consisted of two-dimensional objects. It has been suggested that participants may use different strategies when solving complex tasks of mental rotation or mental folding (Bethell-Fox & Shepard, 1988; Kyllonen et al., 1984). Tasks involving three-dimensional objects are often more difficult, so we used two-dimensional stimuli to make sure that children would not use other strategies but mental manipulation.

Finally, we are interested in the gender differences in mental rotation and mental folding tasks. The goal is to find out whether the gender has an effect on the performance in one or both tasks. The previous studies claim that mental rotation shows sex differences in favor of males starting from the age of 10 (Linn & Petersen, 1985), and some researchers have even suggested that males perform better already at the age of 4 (Levine et al., 1999). Therefore, in the current study we examine whether we can find gender differences in 7- and 8-year-old children. By contrast, gender differences have not been found in mental folding (Voyer et al., 1995), but the reason can be the small number of studies or the complex tasks used to measure the processes. In the present study, we want to examine if we can find differing results.

2. METHODS

2.1. Participants

One hundred twenty-eight children from the first and second grades of the University of Jyväskylä Teacher Training School participated in the study. Ten children were excluded from the study because they did not get consent for participation from their guardians. The remaining 118 children were 89-115 months old. There were 56 boys (M = 100.05, SD = 7.52) and 62 girls (M = 100.48, SD = 6.91). There was no age difference between girls and boys ($t(116) = -0{,}324$, $p=.747$). Children participated to both the computer-based and the paper-and-pencil tests. Because of absence, three children did not participate in the paper-and-pencil tests, whereas four children did not take the computer-based tests. The data was collected during May 2015.

2.2. Materials and procedure

2.2.1. Computer-based tests

The visuospatial skills were measured using the first pilot version of computer-based tests that were developed in Niilo Mäki Institute and realized by Movya Oy. The test battery consisted of nine visuospatial tasks. Tasks assessed visual search, mental transformations, visuospatial memory, estimation of movement and orientation, and impossible figures. Tasks were designed to resemble a computer game and they were optimized for iPad.

The study was conducted at school during the school hours. Tests were done during class under a supervision of the children’s own class teachers. Every child had its own user name and password, which they used to log into the system using iPads. Teachers explained tasks briefly before participants started the tests. In addition, before each task there were illustrated instructions how to perform the task. The sequence of tasks was fixed and participants were able to continue to the next task only after they had finished the previous task. Participants did not get any feedback of their performance.

2.2.1.1. Visual search tasks

Bookshelf. Object finding was assessed with the bookshelf task. The child was first shown an image of a toy for three seconds. After a two-second ISI (interstimulus interval), a square-shaped bookshelf with 3x3 shelving units was shown. The child was instructed to point at the shelving unit where the toy was located. There were four practice items and 18 test items. The number of toys in the bookshelf increased gradually from 3 to 54. In the first ten items the target toy was fully visible and in the last eight items it was partly hidden behind other toys. The child had 15 seconds time to respond. An image of a timer appeared in the right-hand corner of the screen for five seconds if the child had not responded within 10 seconds. Both response time and accuracy of the response were recorded.

Gallery. The Gallery task was used to assess the object recognition skills. The child was shown an imaginary gallery with a square-shaped poster made of a real life photo on the wall. The poster was divided into 3x3 squares. A hand with a photo showing a part of the poster appeared on the screen. The child was instructed to point at the part of the poster where that photo was taken from. There were four practice items with a 3x3 grid and 18 test items without the grid. The child had 15 seconds time to respond. An image of a timer appeared in the right-hand corner of the screen for five seconds if the child had not responded within 10 seconds. Both response time and accuracy of the response were recorded.

2.2.1.2. Mental transformation tasks

Mental Rotation. In this 2D mental rotation task participants had to decide whether two objects were the same or mirror images. Two images were presented side by side, and the child was asked to respond with a button press whether they are the same (=) or different (≠).

Two types of stimuli were used. In the first part of the test, the stimuli were geometrical figures close to the letter F and in the second part they were hands. In the geometrical figure task four different rotations were used (0, 90, 180, 270 degrees clockwise from the upright). All combinations were presented twice (direction of the F x rotation), making a total of 32 items. In the hand task, the same rotations (0, 90, 180, 270) were used but the hands were shown in both palm up and palm down positions, which increased the total amount of items to 64 (hand x rotation x palm up/down).

There were four practice items in both tasks. The presentation order was randomized within the two types of rotation tasks. The child had 15 seconds time to respond. An image of a timer appeared in the right-hand corner of the screen for five seconds if the child had not responded within 10 seconds. Both response time and accuracy of the response were recorded.

MFTC, Mental Folding task for Children. Mental folding task required imagining and choosing how a paper will look like after folding. In the task the child was shown a paper which had dash lines and arrows indicating how it should be folded. The paper varied in shape and size in different items. There were four options of differently folded paper from which the child was instructed to choose the one matching with the correct folding. There were two practice items and 14 task items. Two additional items were added to the end of the original task but they were not used when calculating the scores. The child had 15 seconds time to respond. An image of a timer appeared in the right-hand corner of the screen for five seconds if the child had not responded within 10 seconds. Both response time and accuracy of the response were recorded.

This task is a variant of the MFTC test of Harris et al. (2013). The task uses the same original stimuli with the exception that the outlines indicating the original position of the folded paper have been removed. This was done to prevent a child from using a geometrical symmetry -based solution strategy and to provoke him/her to use mental transformation.

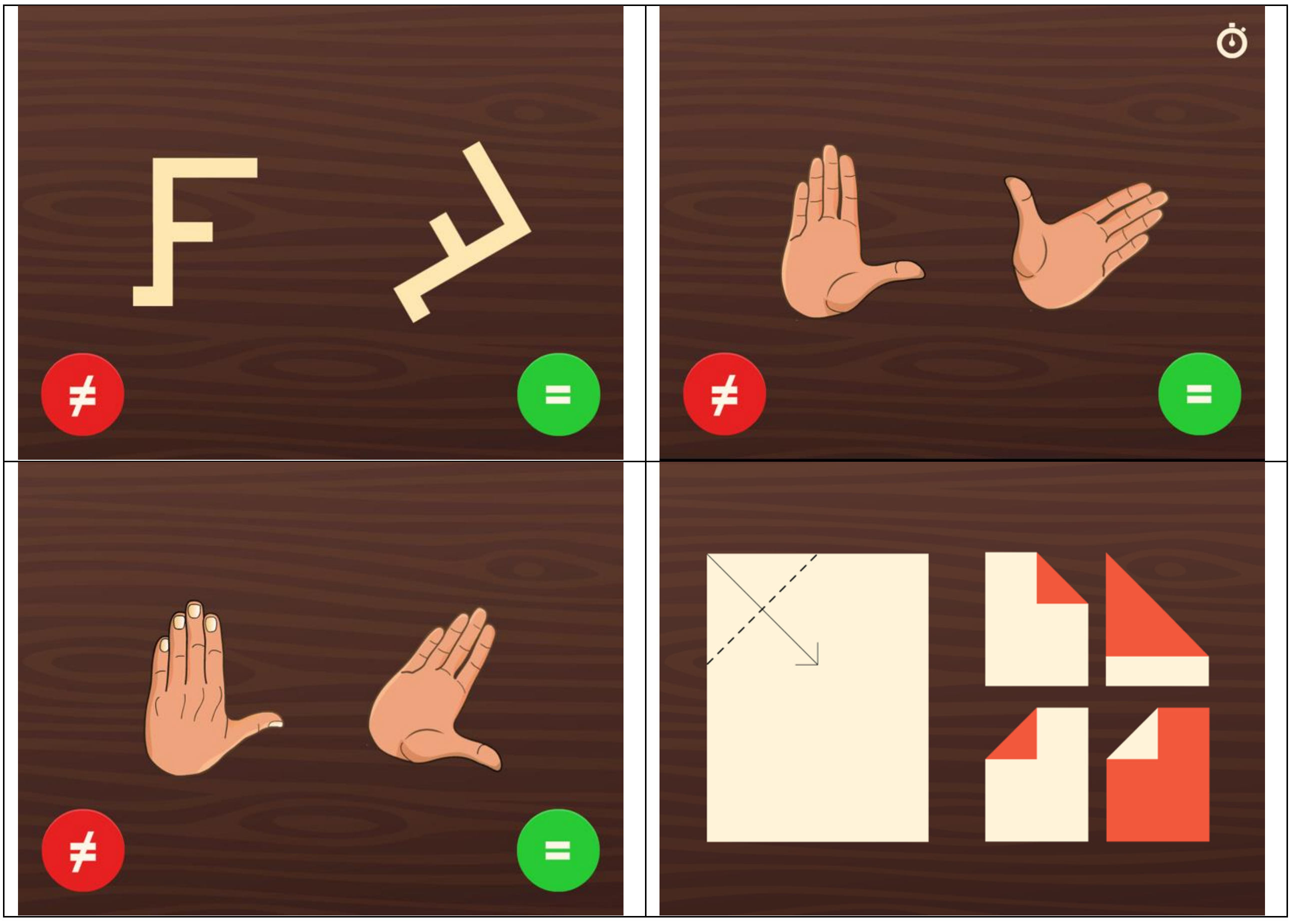

Figure 1. *Screen captures of the three mental rotation task variants (2D object rotation, hand rotation with both hands palms up or down and hand rotation with one palm up and the other down) and the Mental Folding task for Children.*

2.2.1.3. Visuospatial memory tasks

Wooden puzzle. Wooden puzzle task was used to assess object-location memory. The task illustrated a wooden puzzle with 5x5 holes where to place pieces with drawn animal faces. The child was shown a number of pieces placed in the puzzle after which the puzzle was emptied for three seconds. After that, all of the pieces except one were returned to the same positions. The missing piece and three new pieces were shown on the right side of the puzzle. The child was instructed to drag the missing piece back to the same hole where it was. The number of pieces increased stepwise from two to six after four presentations of each number of pieces. Likewise, the presentation time increased every time by a second from five to eight seconds. The child had 15 seconds time to respond. An image of a timer appeared in the right-hand corner of the screen for five seconds if the child had not responded within 10 seconds. If the child failed to respond correctly (correct piece or position) to two of four items with the same number of pieces, the task ended. There were four practice items and the total of 20 task items. The response time and both the correctness of the image and the correctness of the position were recorded.

GrooveBox. The GrooveBox is a variant of the standard Corsi blocks working memory task (Milner, 1971). The child was shown an image of a groovebox with a 5x5 grid of push buttons. Above the square-shaped push buttons there were two small signal lights (red=wait, green=go). While the wait-signal was on, the child was shown a series of lights and after the go-signal turned on, he/she was expected to repeat the series in the same order. The number of lights increased gradually from one to eight. The stimuli were presented one at a time for one second (1000 ms) with 600 ms ISI (interstimulus interval). The item continued until the child pressed the correct amount of lights or the maximum time limit was reached. The maximum response time was 15 seconds for the first two quantities (1,2) and after that an additional four seconds was given for each additional number of lights. An image of a timer appeared in the right-hand corner of the screen five seconds before the end of the response time. There were four practice items. The child was presented test items until s/he made three consecutive errors. The score was the number of correct items and the highest number of correctly remembered lights.

2.2.1.4. Estimation of movement and orientation tasks

CakeMachine. The cake machine task was used to assess the child’s ability to evaluate movement. In this task a layer cake rolled into an opaque cake machine and the child had to estimate when the cake has moved into the indicated place by pressing a button to spurt a garnish in the middle of the cake. When the cake came out of the machine, the child saw where the whipped cream garnish had landed. Three speeds (300, 400, 500 px/s), two directions (left-right, right-left) and three lengths of the machine (210, 410, 610) were used to vary the time and length of the movement estimation. There were six practice items with diminishing transparency of the machine and 54 task items presented randomly within three blocks. The score was the time difference between the child’s responses and the precise responses.

Map Navigation task (The Parachutist). The Parachutist task measured orientation and navigation skills. The tablet screen was divided into two images. On the left side there was a 3D image of buildings of different sizes and shapes, and on the right side there was a 2D overview map of these buildings. A parachutist landed on the roof of one of the buildings and the participant's task was to point the same building on the map. The task difficulty varied depending on the number of buildings in the task (from 3 to 8) and the amount of rotation of the map from the landscape (0–180). The task consisted of 3 practice items and 16 items.

2.2.1.5. Impossible figures task

The Impossible figures task requires figure judgment. The participants saw 2D drawings of figures, and they had to indicate by means of a button press whether the figures were "possible" or "impossible". Participants were instructed that "possible" figures looked as if they were a real 3D object that you could reach out, grab and hold, whereas it was "impossible" for an impossible figure to exist in real life because one or more of the lines or corners were out of place. The stimuli were taken from the original set of possible / impossible stimuli (Schacter, Cooper, & Delaney, 1990). From the original 168 items in the Schacter list, 58 items were selected to this task. The items were selected from two sources: 48 items were picked from the Carrasco & Seamon (1999, see their table 4) list of equal and extreme items. The list of items was complemented with items from Chan (2010, see his tables 2 and 3), who made a Rasch analysis on the difficulties of the items and based on that made a short battery of 18 items. Ten of these items were different from the Carrasco and Seamon list, and were added to our task. Therefore, the final task consisted of 6 practice items and 58 randomly presented items (30 possible and 28 impossible). The items were redrawn to vector images from the original bitmap drawings received from Anja Soldan (John Hopkins Medical University, US) with the permission of original author D. Schacher.

2.2.2. Paper-and-pencil tests

In addition to computer-based visuospatial tests, participants did paper-and-pencil tests that assessed their reading and math, non-verbal reasoning, spatial visualization and visual perception skills. The tests were done in a group situation in classrooms.

Technical reading. Technical reading skills were assessed using two subtests, Find Words and Nonsense Words, from the Word Chain test (Nevala & Lyytinen, 2000). The first task consisted of ten word chains, including four to six words written together. Children were asked to recognize the words and separate them by drawing vertical lines between words. The time limit for the task was one minute and twenty-five seconds. One point was given for each correctly separated word (maximum score: 40). The second task required children to recognize pseudo words from meaningful words. The task presented 25 words and 25 pseudo words, and participants were instructed to draw a vertical line across the word if the word was a pseudo word. The time limit for the task was 50 seconds. Children received one point for each correct answer (maximum score: 25). Cronbach's alpha coefficients were .72 for the first task and .86 for the second task (Nevala & Lyytinen, 2000).

Reading fluency. Reading fluency was assessed using LUKSU test (Suokas, 2009). In the test, children read sentences and decided if they were sensible or not. Children were given three minutes to complete 70 items. The score was the number of correct answers (maximum score: 70). Even though the test requires the child to understand what s/he is reading, the sentences are easy, so the test is more a measurement of reading fluency. Cronbach's alpha coefficient was .95.

Arithmetic competence. Arithmetical skills were measured using the Basic Arithmetic test, BAS (Räsänen & Aunola, 2007). In the test, a maximum of 28 items containing 14 addition items (e.g., 2+1=; 3+4+6=) and 14 subtraction items (e.g., 4–1=; 20−2−4=) could be attempted within a 3-minute time limit. Task difficulty increased gradually across the test. The score was the number of correct answers (maximum score: 28). The test is a measurement of arithmetic knowledge that can be counted as a combination of speed and accuracy of performance. It represents the cognitive domain of knowing in mathematical thinking (Mullis & Martin, 2013). Cronbach's alpha coefficient was .92.

Verbal problems. The verbal problem task consisted of 16 simple verbal mathematical problems (e.g., "Sami has seven candies and he gets three more. How many candies does he have now?") Task difficulty increased across the test. The score was the number of correct answers given within a four-minute time limit (maximum score: 16). Solving verbal problem tasks requires applying arithmetic knowledge. Cronbach's alpha coefficient was .78.

Arithmetic reasoning. Children did a task of arithmetic reasoning, which was adopted from the Arithmetic Reasoning Test (Räsänen, 2000). The test required the child to continue a series of three numbers (e.g., 3, 5, 7) by adding a fourth number that best fit the series (e.g., 9). For each question, there were four response alternatives. Children were given four minutes to complete 16 items. Every correct answer gave one point, so the maximum score was 16. The test measures arithmetic reasoning skills, whereas the other mathematical tasks assess more the arithmetic knowledge. Cronbach's alpha coefficient was .89.

Non-verbal reasoning. Non-verbal reasoning was assessed with the subtest of Matrix Reasoning from the Wechsler Intelligence Scale for Children (WISC-IV) (Wechsler, 2003). In the test, the child was instructed to complete the missing portion of a picture matrix by choosing a correct pattern from a set of five response options displayed below the matrix. The test falls into the perceptual reasoning index, and it measures fluid reasoning, i.e. the child's ability to comprehend nonverbal shapes and designs and complete or correct the missing or incorrect aspects of those designs. It requires the comparison and classification of objects and understanding their details. The test consisted of 35 items, and the children got 8 minutes 45 seconds to complete it. Each correct answer gave one point, so the maximum score was 35. Cronbach's alpha coefficient was .85.

Spatial visualization. Spatial visualization was measured with the subtest of Spatial Relations from the Woodcock and Johnson test battery (Woodcock, 1997). The test requires identifying the subset of pieces needed to form a complete shape, with multiple-point scored items (i.e., "Two of these pieces (a, b, c, d) go together to make this (e). Tell me which two pieces."). Spatial relations task involves detecting spatial forms or shapes and rotating or manipulating them mentally. A maximum of 31 items could be attempted within a three-minute time limit (maximum score: 31). Cronbach's alpha coefficient was .86.

Visual perception. Visual perception skills were measured using the subtests of Figure-Ground and Spatial Relationships from the Test of Visual-Perceptual Skills (TVPS) (Martin, 2006). In the Figure-Ground test, the child was asked to find one design among many within a complex background. In the Spatial Relationships test, the child was shown a series of designs on a page and then asked to choose the one that is different from the rest; it could differ in a detail or in the rotation of all or one part of the design. For both subtests, a total of 16 items could be attempted within a 4-min time limit. Each correct answer was scored as 1, so the maximum score was 16. In the current study, the two subtasks were selected to assess visual perception since they load on different factors (Martin, 2006). The Spatial Relationships task assesses the basic processes of visual perception, i.e. the child's ability to perceive the positions of objects in relation to oneself and/or other objects. By contrast, the Figure-Ground task assesses the child's ability to identify an object from a complex background. It assesses the complex processes that are the last perceptual tasks to develop. Cronbach's alpha coefficients were .70 for the Figure-Ground task and .78 for the Spatial Relationships task.

2.3. Statistical Analyses

In the study we are interested in mental transformations, so the computer-based subtests Mental Folding and Mental Rotation were chosen to the analyses. The variables we examined in the mental folding task were response time, accuracy (the number of correct responses) and folding efficiency that was computed by dividing median time by the percentage of correct answers. We excluded the responses that were given faster than in 400 milliseconds because it was estimated that in shorter time it was not possible to decide between four options without guessing. By contrast, in mental rotation task we examined response time, accuracy (the number of correct responses) and linearity. We computed separate median times for each rotation angle and then compared the means and standard deviations of the median times. We excluded responses if the response time was less than 200 milliseconds or more than three times standard deviation. Mental rotation task was divided into three different parts depending on the complexity: 1. Rotation of objects, 2. Rotation of hands: both hands were presented palm/back up and 3. Rotation of hands: one hand was presented palm up and other back up (see Figure 1).

The analyses were made with IBM SPSS Statistics 24. First we conducted the reliability analyses to examine the reliability of the tasks. Next an independent samples t-test was performed to find out whether there are significant differences in the results depending on the grade. We also examined correlations between the tasks with Pearson correlation coefficients and after that, two regression analyses were conducted to find out if the mental rotation skills can predict the math score. Finally, we performed an independent samples t-test to examine the gender differences.

3. RESULTS

3.1. Reliability

First we wanted to examine the reliability of the mental rotation and mental folding tasks. The separate reliabilities were computed for three different rotation tasks using Split half method. In every task we computed reliability using both the correctness and the response time of items. In the object task, Cronbach's alphas for the correctness and for the response time were .78 and .69. The percentage of correct answers was 58.5 %. For the hands presented same side up, Cronbach's alphas for the correctness and the response time were .67 and .74. Correctness was 54.4 %. For the hands presented different side up, Cronbach's alphas for the correctness and for the response time were .55 and .76. Correctness was 50.1 %.

In the mental folding task, the 14 original items were included in the analyses ($\alpha = .67$). The additional more complex items 15 and 16 did not improve the reliability, so they were excluded. The percentage of correct answers was 46.3 %. The correctness and the correlations of items are presented in Table 1.

Table 1. *The Correctness and Corrected Item-Total Correlations of the Mental Folding items.*

Item	Correctness	Correlation
1	0,45	0,043
2	0,34	0,06
3	0,71	0,484
4	0,32	0,235
5	0,64	0,348
6	0,66	0,583
7	0,62	0,56
8	0,54	0,478
9	0,38	0,351
10	0,22	0,031
11	0,55	0,246
12	0,51	0,278
13	0,28	0,27
14	0,27	0,036

3.2. Descriptive Statistics

Table 2 presents the descriptive statistics of the paper-and pencil test, and the differences in the scores between first- and second-graders. We performed independent samples t-test to find out the differences. The second-graders had significantly better scores in the reading tasks Nonsense Words and LUKSU, mathematical tasks BAS and Verbal problems, as well as in Woodcock-Johnson Spatial Relations. The third mathematical task Arithmetic reasoning was not normally distributed, with skewness of -1.11 (SE = 0.23), which shows that the task was too easy for the children of this age.

Table 3 presents the descriptive statistics for mental folding and mental rotation tasks, and the differences between the two grades, which were also examined with independent samples t-test. The second-graders were significantly faster to find a correct answer in all three rotation tasks. However, there was no difference in the scores between the two grades. The mental folding task did not show any significant differences.

The linearity of response times in mental rotation task was examined with the task of object rotation. The median times were computed for each rotation angle. Median0 (M = 1457.54, SD = 624.41), Median90 (M = 1643.52, SD = 961.83), Median270 (M = 1493.54, SD = 874.70) and Median180 (M = 1773.70, SD = 1071.74) did not differ, so there was no linearity of response times depending on the rotation angle.

Table 2. *Means (M) and Standard Deviations (SD) of the paper-and-pencil tests, and differences between Grades 1 and 2.*

		Grade 1		Grade 2			
Task	Range	M	SD	M	SD	F	p
Find Words	0-40	8,67	5,94	10,72	5,14	-1,974	.051
Nonsense Words	0-25	9,74	5,06	11,93	6,06	-2,094	.039
LUKSU	0-70	21,36	8,82	31,07	10,56	-5,354	.000
BAS	0-28	11,17	5,14	16,21	4,69	-5,486	.000
Verbal Problems	0-16	5,64	2,28	7,49	2,03	-4,599	.000
Arithmetic Reasoning	0–16	10,43	4,08	11,68	3,11	-1,849	.067
TVPS Figure-Ground	0-16	8,03	2,57	8,46	2,35	-0,918	.361
TVPS Spatial Relationships	0-16	10,97	2,39	11,56	2,04	-1,438	.153
W-J Spatial Relations	0-31	14,91	2,57	15,82	1,91	-2,154	.033
WISC Matrix Reasoning	0-35	19,76	4,38	20,39	3,95	-0,806	.422

Table 3. *Means (M) and Standard Deviations (SD) of Mental Folding and Mental Rotation tasks, and differences between Grades 1 and 2.*

	Grade 1		Grade 2			
Task	M	SD	M	SD	F	p
Folding Score	6,17	2,90	6,71	3,02	0,929	.337
Folding Median	4844,25	1719,09	4949,51	2084,56	0,086	.770
Folding Efficiency	125,84	71,85	109,76	56,40	1,612	.207
R1 Score	17,39	5,17	17,44	6,07	0,002	.965
R1 Median	1634,70	811,70	1357,24	561,17	4,445	.037
R2 Score	15,39	4,95	15,76	5,24	0,153	.696
R2 Median	2118,77	1596,75	1541,07	780,09	5,887	.017
R3 Score	14,53	4,58	14,09	4,72	0,249	.619
R3 Median	2320,83	1829,99	1597,86	1060,82	6,535	.012

Note: R1 = Rotation of objects, R2 = Rotation of hands, same side; R3 = Rotation of hands, palm and back.

3.3. Correlations

We examined correlations with Pearson correlation coefficients. Table 4 shows correlations among paper-and-pencil tests. All correlations were positive. All the paper-and-pencil tests correlated significantly with each other, with the exception of TVPS Figure-Ground and TVPS Spatial relationships that correlated only with part of the tests. The subtest Figure-Ground correlated significantly with TVPS Spatial relationships, Woodcock-Johnson Spatial relations, WISC Matrix Reasoning, LUKSU and Arithmetic Reasoning. The other subtest Spatial Relationships had a significant correlation with all the other tests apart from the reading tests Find Words and Nonsense Words.

Table 5 shows the correlations among variables in the mental folding and mental rotation tasks. Folding Score (the number of corrects responses) and Folding Median (the median response time) had a significant positive correlation: it took more time to solve the task if children gave the right answer. In addition, Folding Score correlated positively with two rotation scores: rotation of objects (R1) and same side up presented hands (R2). Folding Median correlated significantly with the R1 score and the R2 median: the children who used more time to solve the mental folding task, needed more time also in the hand rotation task, and they were more accurate in the object rotation task. In addition, Folding Efficiency correlated positively with Folding Score and negatively with Folding Median: the children who were more efficient, solved mental folding task more accurately but more slowly. Furthermore, the median times of all rotation tasks correlated positively with each other and with the R1 score, and the median times of both hand rotation tasks had a correlation with the R2 score. That indicates that the children who used more time solving the rotation tasks, had more accurate responses in one or two rotation tasks. In addition, the R1 score correlated positively with all the variables apart from the folding efficiency and the R3 score. It would seem that the children who were talented in object rotation, used more time to solve all the tasks, and they were skilled also in other tasks, apart from the most complex hand task. In fact, R3 score correlated only with R2 score: the children who got better scores in one hand task, got better scores in other as well.

Table 6 presents how mental folding and rotation tasks correlate with paper-and-pencil tests. Folding score correlated positively with WISC Matrix Reasoning, Woodcock-Johnson Spatial relations and TVPS Figure-Ground. The children who had good scores in mental folding, were more accurate in these paper-and-pencil tests as well. In addition, Folding Efficiency correlated negatively with Woodcock-Johnson Spatial Relations and Find Words. It seems that the children who were more efficient in mental folding, performed also well in these two paper-and-pencil tasks.

On the contrary, the R1 and R2 score correlated positively with WISC Matrix Reasoning, TVPS Spatial Relationships, Arithmetic Reasoning and Verbal Problems. It appears that the children who were skilled in mental rotation task, had good scores also in these paper-and-pencil tests. In addition, the R1 score correlated positively with LUKSU, so the children who had good object rotation scores, were skilled in reading test LUKSU as well. The only median time that correlated with paper-and-pencil tests was the R3 median, which correlated positively with WISC Matrix Reasoning. Therefore, it seems that if children took more time to solve the most complex hand task, they had better scores in Matrix Reasoning.

Table 4. *Correlations among paper-and-pencil tests.*

	1	2	3	4	5	6	7	8	9	10
1. Find Words	–									
2. Nonsense Words	.67**	–								
3. LUKSU	.58**	.72**	–							
4. BAS	.54**	.47**	.66**	–						
5. Verbal Problems	.54**	.59**	.68**	.66**	–					
6. Arithmetic Reasoning	.32**	.37**	.47**	.52**	.54**	–				
7. TVPS Figure-Ground	.12	.10	.22*	.15	.18	.22*	–			
8. TVPS Spatial Relationships	.18	.18	.41**	.27**	.32**	.32**	.33**	–		
9. W-J Spatial Relations	.30**	.23*	.41**	.42**	.34**	.35**	.20*	.44**	–	
10. WISC Matrix Reasoning	.43**	.30**	.41**	.42**	.48**	.38**	.36**	.51**	.41**	–

*p < .05. **p < .001, two-tailed.

Table 5. *Correlations among the variables of Mental folding and Mental rotation tasks.*

	1	2	3	4	5	6	7	8	9
1. Folding Score	–								
2. Folding Median	.37**	–							
3. Folding Efficiency	-.55**	.36**	–						
4. R1 Score	.34**	.29**	-.04	–					
5. R1 Median	.18	.14	-.03	.47**	–				
6. R2 Score	.29**	.17	-.11	.39**	.15	–			
7. R2 Median	.18	.25**	.01	.33**	.43**	.27**	–		
8. R3 Score	.01	.06	-.06	.07	.14	.40**	.18	–	
9. R3 Median	.14	.17	.02	.29**	41**	.33**	.89**	.14	–

Note: R1 = Rotation of objects, R2 = Rotation of hands, same side; R3 = Rotation of hands, palm and back.
*p < .05. **p < .001, two-tailed.

Table 6. *Mental Folding and Mental Rotation Correlations with paper-and-pencil tests.*

	Find Words	Nonsense Words	LUKSU	BAS	Verbal Problems	Arithmetic Reasoning	TVPS Figure-Ground	TVPS Spatial Relationships	W-J Spatial Relations	WISC Matrix Reasoning
Folding Score	.06	-.07	.03	.10	.04	.07	.21*	.06	.24*	.34**
Folding Median	-.14	-.18	-.08	-.01	.05	.11	.14	.15	.07	.14
Folding Efficiency	-.21*	-.19	-.16	-.15	-.07	-.05	-.07	.08	-.27**	-.15
R1 Score	.05	.07	.25**	.08	.19*	.21*	.17	.35**	.17	.39**
R2 Median	-.10	-.05	.00	-.08	-.05	.00	.05	.03	-.05	.11
R2 Score	-.03	.06	.15	.18	.19*	.19*	.14	.28**	.18	.28**
R2 Median	-.07	-.09	-.04	-.06	.00	.00	.11	.02	.02	.14
R3 Score	.03	-.04	.03	.04	.04	.12	.08	.16	.12	.05
R3 Median	-.03	-.04	-.05	-.03	.03	-.01	.11	.03	.07	.19*

Note: R1 = Rotation of objects, R2 = Rotation of hands, same side; R3 = Rotation of hands, palm and back.
*p < .05. **p < .001., two-tailed.

3.4. Regression analysis

Two linear regression analyses were conducted to test whether mental rotation skills can predict mathematical skills. The mental folding test was not included in the analyses because it did not have a correlation with math. Mathematical skills were examined separately based on the cognitive domain. However, Basic Arithmetic Test was not included since it did not correlate with mental rotation. To control that the explained variance between the rotation score and the mathematical skills would not result from more general factors, like overall good skills in performing academic and other tasks or more general visuo-spatial reasoning skills (non-verbal IQ), the mental rotation score was chosen for a predictor together with a reading test LUKSU and a non-verbal reasoning test WISC Matrix Reasoning. We made two regression analyses where the dependent variables were Verbal Problems and Arithmetic Reasoning. We used a stepwise method where the independent variables were firstly the score of mental rotation of objects, secondly the score of rotation and LUKSU, and thirdly the score of rotation and Matrix Reasoning.

Mental rotation was a significant predictor of Verbal Problems test ($F (1, 109) = 4{,}082$, $p<.05$). The children who performed well in mental rotation had good scores in Verbal Problems as well. However, mental rotation predicted only 3% of the variance ($R^2 = .03$). When we added LUKSU to the independent variable list, the model predicted the Verbal Problem score significantly ($R^2 = .47$, $F (1,109) = 98{,}996$, $p< .001$). However, only LUKSU had main effect ($\beta = .69$, $p <. 001$): the children who had good scores in LUKSU, had good scores also in Verbal Problems task. Therefore, mental rotation was not a significant predictor. In third model, when independent variables were rotation score and Matrix Reasoning, the model predicted Verbal Problems significantly ($R^2 = .215$, $F (1,109) = 31{,}172$, $p<.001$), but only Matrix Reasoning was a significant predictor ($\beta = .47$, $p<.001$). The children who were skilled in WISC Matrix Reasoning, had good scores in Verbal Problem task as well.

In the second analysis, mental rotation predicted significantly the score of Arithmetic Reasoning Test ($F (1,109) = 4{,}762$, $p<.05$). The children who had good scores in the mental rotation task, were skilled also in Arithmetic Reasoning test, but the coefficient of determination was modest ($R^2 = .03$). When adding LUKSU to the list of independent variables, the model predicted Arithmetic Reasoning ($R^2 = .22$, $F (1,109) = 32{,}215$, $p<.001$) but only LUKSU was significant predictor ($\beta = .48$, $p<.001$). Therefore, the children who performed well in LUKSU, had good scores also in Arithmetic Reasoning test. Similarly, when independent variables were rotation score and WISC, the model was a good predictor ($R^2 = .13$, $F (1,109) = 16{,}666$, $p<.001$) but only WISC had main effect ($\beta = .36$, $p<.001$). The children who performed well in WISC Matrix Reasoning, had good scores in Arithmetic Reasoning test as well. In conclusion, when the effects of the more general factors were controlled, at this age the mental rotation task did not add the predictive power of the regression model to mathematical skills.

3.5. Sex differences

Finally, we performed an independent samples t-test to find out, if we can find differences in the scores between boys and girls. In the Basic Arithmetic test, boys (M = 14.68, SD = 6.17) performed better than girls (M = 12.57, SD = 4.88), t (109) = 4,230, $p<.05$. On the contrary, girls (M = 8.91, SD = 2.41) got better scores than boys (M = 7.42, SD = 2.32) in TVPS Figure-Ground (t (109) = 11,089, p=.001). In addition, girls (M = 15.93, SD = 1.97) performed better than boys (M = 14.87, SD = 2.33) also in Woodcock-Johnson Spatial Relations task (t (109) = 6,772, $p<.05$). Other tasks did not show significant gender differences.

In the mental folding and mental rotation tasks, the results show significant differences in the mental folding score, where girls (M = 7.23, SD = 2.80) got better scores than boys (M = 5.53, SD = 2.89), t (111) = 10,113, $p<.01$. That means that girls chose the correct answer more often than boys. In addition, folding efficiency showed gender differences: girls (M = 98.24, SD = 43.89) were more efficient solving mental folding items than boys (M = 141.37, SD = 79.85), t (107) = 12,702, p=.001. However, the response time of mental folding task did not differ between the genders, which proves that boys and girls were equally fast choosing the right answer. Mental rotation did not show significant sex differences, although in the mental rotation of objects, girls (M = 18.32, SD = 5.27) had higher mean score than boys (M = 16.41, SD = 5.82), t (112) = 3,377, p=.07.

4. DISCUSSION

The main goal of the present study was to increase the knowledge of the processes of mental folding and mental rotation in 7- and 8-year-old children. First we examined the reliability of the tests. The well-known acceptable level for coefficient Alpha is .70 (Nunnally, 1978). Some of the tasks did not attain that level; however, the reliability could have suffered from measurement error. The reliability of mental folding task was lower than .70, so we wanted to examine individual items more closely. As we can see in Table 1, the order of the tasks is not becoming more difficult in a logical order. The most difficult items were 10, 14, 13, 4 and 2, and the easiest was the item 3. If the items became more difficult in a logical order, perhaps it would be easier for the children to understand and be more motivated to complete the task. We can see that for some items the corrected item-total correlations were very weak. It would seem that especially items 1, 2, 10 and 14 correlate poorly with other items. Almost all of them are also among the most difficult items. The internal reliability was lower than in the original test (Harris et al., 2013). However, this computerized version differed in one significant factor from the original paper-and-pencil version. In the original version, the outlines of the paper were visible and marked with a dotted line. However, keeping the outlines changes the task into a symmetry judgement instead of a folding task (personal communication prof. Andrea Frick, 11.2.2015). Leaving the outlines out kept the task as a folding task, but at the same time made it much more difficult. It could be that this lowered the reliability. However, it might be reasonable to examine the test items and their order more closely, and possibly modify the order of presentation. Likewise, an interesting study would be to compare the performances in the original version (symmetry judgement) and our version (folding task), while keeping all other factors in the task the same.

The tasks that required rotation of hands were too difficult for children at this age. The reliabilities of the tasks were low, probably because most of the children guessed their responses. In addition, the response times of the rotation tasks did not show linearity depending on the rotation angle, which has been taken as a sign of mental rotation (Cooper & Shepard, 1973). It seems that most of the children did not mentally rotate images. The results are consistent with the early view of Piaget and Inhelder (1971) who claimed that children are only starting to mentally rotate images at the age of 7 or 8. Some studies have found mental rotation in younger children as well (Frick et al., 2013), but in these studies the stimuli have been more simple. It seems that the tasks we used are too complex for young children to solve using a mental rotation strategy. The earlier studies have shown that young children benefit from motor activities more than older children and adults when they solve mental transformation tasks (Frick et al., 2009). The findings are consistent with the observations in the current study: in the hand task many children rotated their own hands trying to solve the task.

One question of interest was whether there are differences in the results among first- and second-graders. As we could assume, the differences were significant in favor of second-graders in almost all the reading and mathematical tasks. Reading and mathematics are subjects that are studied at school, so second-graders normally perform better in these tasks. We did not get significant results in the arithmetical reasoning task, but that is probably because the task was too easy and did not sort out the children well enough. In the mental folding task, the median time and the score did not differ between the grades. That proves that there are no differences in the skills of mental folding between grades 1 and 2. That supports the view of Harris et al. (2013), who suggested that some mental folding skills appear already before the first grade. The scores did not differ either in the mental rotation tasks, but in all the rotation tasks first-graders rotated items significantly slower. It seems that first-graders are as good as second-graders to conclude the right answer in mental rotation tasks, but second-graders are faster to choose the correct answer. The results are similar than Kail et al. (1980) suggested when they found out that mental rotation accelerates with age, although most of the children did not mentally rotate the images.

4.1. Relations to other tests

Almost all paper-and-pencil tests correlated with each other, which shows that there are similar processes behind reading, mathematical and visuospatial tasks. WISC Matrix reasoning, which is a measurement of non-verbal reasoning, had a very significant correlation with all the other tasks, so it seems that all visuospatial tasks benefit from non-verbal reasoning skills. Similarly, the scores of mental folding and mental rotation correlated with WISC Matrix reasoning, so also the mental transformations seem to have a connection with non-verbal reasoning.

Two rotation scores correlated with the folding score, so it would seem that the processes behind both tasks are at least partly similar, and that solving the tasks requires some similar skills. The processes cannot be completely similar though, since the tasks correlated differently with the paper-and-pencil tests. Mental rotation but not mental folding had a connection with TVPS Spatial Relationships. The Spatial Relationships task assesses the basic processes of visual perception, i.e. the child's ability to perceive the positions of objects in relation to oneself and/or other objects, like rotation. Therefore, the correlation indicates that the current mental rotation task measured the same processes, so some children probably used a mental rotation strategy. However, the mental rotation task did not have a relation with Woodcock-Johnson Spatial Relations, which was a surprising finding. Spatial Relations measures the processes of mental manipulation, i.e. modifying or rotating objects mentally that are the basic processes behind mental rotation as well. The lack of correlation indicates that most of the children at this age did not use mental rotation to solve the rotation tasks. It supports the finding that linearity of response times was not found. In contrast, the mental folding task correlated with Woodcock-Johnson Spatial Relations, so it appears that children might have used mental manipulation to solve the task. Furthermore, the mental folding score correlated with TVPS Figure-Ground that assesses the ability to identify an object from a complex background. The relation indicates that there can be some similar processes behind mental folding. However, the correlation was not very strong so we need more research on the topic to understand better the processes behind mental folding.

Mental rotation had a correlation with two mathematical tests that measure applying and reasoning. Therefore, it would seem that there is a connection between mental rotation and more complex mathematical thinking that does not include knowing facts. Earlier studies have not been interested in studying this correlation in children. Yet, the results are in line with other research findings with adults that have shown that the score of mental rotation has a relation to math scores (Casey et al., 1995). However, the correlation was not strong, so more research is needed to understand better this relation in children. Mental folding had no connection with math, which supports the view that the processes behind mental rotation and mental folding are different. In addition, mental rotation but not mental folding score had a connection with the reading test LUKSU. The children who were skilled in the object rotation task, performed well also in LUKSU. The correlation can probably be explained with non-verbal reasoning skills that had a strong correlation with both of the tasks.

We assumed that mental rotation of hands would correlate more with complex mathematical task that requires reasoning. The connection was found only with the easier hand task where hands were presented same side up, but even that correlation was not as strong as the correlation between the object rotation and the reasoning task. A relation with the more complex hand task was not found. In fact, the score of the more complex hand task did not correlate with any paper-and-pencil task, not even with Matrix Reasoning that measures non-verbal reasoning skills. The lack of correlation indicates that the task was too difficult for 7- and 8-year-olds, and most of the children had to guess their responses. Because of that, the reliability of the task was quite low, which can be one reason why we did not find correlations between the task and math scores. However, the score of the complex hand rotation task correlated with the score of the other hand task, and the median time correlated with WISC Matrix Reasoning. Thus, the children who had good scores in one hand task, more likely had good scores in other as well. Furthermore, the children who took more time to solve the complex hand task, had better scores in Matrix Reasoning. This indicates that slower responses were less likely guessed, so they correlated with non-verbal reasoning. This gives a reason to conclude that the task might work well with older children.

Mental transformations predicted mathematical skills very poorly. Mental folding did not have any effect, and mental rotation predicted weakly the mathematical tasks that required applying and reasoning. The results showed that at this age reading skills and nonverbal reasoning skills were better predictors of mathematics than mental rotation. Earlier studies have proved that mental rotation can predict math scores in adults (Casey et al., 1995). The current study indicates that at the age of 7 and 8, children's mental rotation skills do not predict well the performance in mathematics. However, it seems that most of the children did not use mental rotation when they solved the rotation tasks, so that can be one reason why we did not find a strong connection. Therefore, the study should be replicated with older children who are already capable to rotate images mentally.

4.2. Sex differences

We were also interested in the gender differences among the tasks. The results show that boys had better scores in the Basic Arithmetic test that is a measurement of arithmetic knowledge. That suggests that at the age of 7 and 8 males perform better in mathematical tasks that require knowing the facts and concepts. The results are inconsistent with a meta-analysis (Hyde, Fennema, & Lamon, 1990), which showed that girls were slightly better at computing at elementary school. Other mathematical tasks that require applying and reasoning did not differ between the genders. On the contrary, girls got better scores in TVPS Figure-Ground task that measures the visuo-perceptual skills needed in identifying an object from a complex background. These are complex processes that are the last perceptual tasks to develop (Martin, 2006), so it may be that girls develop these skills faster than boys. Girls also performed better in Woodcock-Johnson Spatial Relations task that requires mental manipulation and rotation. Girls had a higher mean score in the object rotation task as well, although the difference was not significant. However, the results are inconsistent with the previous findings where boys were better at mentally rotating images already at the age of 4 (Levine et al., 1999). The current study claims that 7- and 8-year-old boys and girls are equally skilled solving mental rotation tasks.

On the contrary, girls had better scores in the mental folding task and they were more efficient solving the task. That indicates that girls are more skilled to fold images mentally at the age of 7 and 8. The previous studies have not found sex differences in mental folding tasks, so the results are noteworthy. We can speculate that the differences might result from sociocultural reasons. The girls might have more experience of these kind of tasks because girls do arts and crafts more often than boys do. The earlier studies have shown that training can improve spatial skills in both children and adults (Uttal et al., 2013), and that playing games develops spatial skills in childhood (Caldera et al., 1999), so it can be that the girls' play preferences develop their mental folding skills. We wanted to examine only the skill of mental manipulation and to eliminate the possible effects of complexity. Therefore, we used two-dimensional items that differ from complex, multi-phased stimuli that researchers have often used for measuring mental folding. That can be one reason why we found divergent results. It would be interesting to replicate the study with different-aged children and examine whether we can find similar results. In addition, it would be interesting to control the hobbies and then examine if the sex differences would disappear.

4.3. Limitations and future directions

Since the study was a pilot and the current computer-based tests were used for the first time, we found out some problems that should be improved for the next version. For some children it might have been hard to understand all tasks, since there were only short illustrated instructions before the tasks. That should be improved for the next test version to make sure that wrong answers are not caused by too ambiguous instructions. In addition, the lack of feedback might have confused some children. Because participants did not get any feedback after their responses, they could have been uncertain if they understood the task correctly. That should also be modified for the next test version. The feedback of correct answers is important to motivate the children to finish the task and not starting to guess the answers. It has been noticed that some of the weak and even more competent children do not give their best performance in computer-based tests compared to paper-and-pencil tests because it is easy to respond by making a quick guess (personal communication Dr. Jarkko Hautala, 1.12.2016). The test could be modified by making it more gamelike; correct answers and fast response time could collect scores, as children are used to in their videogames. In addition, the items could be divided into different levels that become more complicated in a logical way. Thus, if children can solve easier tasks, they would be motivated to move on to the more difficult level. For example, that could be done with hand rotation tasks, where the different sides up presented hands were more difficult than same side up presented hands. In addition, the presentation order of mental folding items should be considered. The levels could help children to concentrate on the task, especially when there are lots of items, and reduce guessed responses. Furthermore, the tasks should be stopped automatically after too many incorrect responses so that children would not have to continue solving too difficult tasks for too long. Besides the problems of the test, there were some problems with technology that could have affected the scores. In addition, because of limited resources, tests were conducted in group situations in classrooms, which made it more difficult to monitor the children to make sure they understood the task and were not guessing their responses.

Some of the test items should also be modified, since they were too difficult for the children of this age. Especially the hand rotation tasks were too difficult. However, the easier hand task correlated with other tasks, which shows that the task might work with older children. More complex hand task did not have a relation to many other tests but it is probably because the reliability was low and most of the children needed to guess the items. This task too could probably be used with older children. The mental transformation tasks did not correlate strongly/at all with the visuospatial tasks that measured mental manipulation. In addition, the linearity of response times was not found in mental rotation task. It indicates that most of the children did not mentally rotate the objects. We were particularly interested in examining mental manipulation behind mental rotation and mental folding processes, so some changes for the test need to be made. It seems that at this age the mental rotation skills should be measured with more simple tests.

In conclusion, we found out more information on the processes behind mental transformations in children. However, it seems that children at the age of 7 and 8 do not mentally rotate images. That is why we did not get information on the strategy of mental rotation, but on the other processes that children used to solve the task. It appears that in childhood the processes of mental rotation and mental folding are partly similar, since they correlated with each other, and partly different because correlations with paper-and-pencil tests differed. The correlation with mathematics was only found in mental rotation, as the meta-analysis of Harris et al. has also shown (2013). However, since the children were probably not using mental rotation, it seems that there are other common processes between mental rotation and mathematics. Mental rotation was not a strong predictor of mathematical skills but more significant results might be found if participants would solve the task using a mental rotation strategy. However, we found remarkable results about the gender differences in mental folding; the girls had significantly better scores. Earlier studies have not demonstrated sex differences in mental folding. Yet, the results should be replicated with a new improved test version, so that the findings would be more reliable.

5. REFERENCES

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles, 20*(5-6), 327-344. doi:http://dx.doi.org/10.1007/BF00287729

Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 14*(1), 12-23. doi:http://dx.doi.org/10.1037/0096-1523.14.1.12

Caldera, Y. M., Mc Culp, A., O'Brien, M., Truglio, R. T., Alvarez, M., & Huston, A. C. (1999). Children's play preferences, construction play with blocks, and visual-spatial skills: Are they related? *International Journal of Behavioral Development, 23*(4), 855-872.

Carrasco, M., & Seamon, J. G. (1999). The effect of study time on priming possible and impossible figures in the object decision test. *Psicothema, 11*(4), 801-813.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31*(4), 697-705. doi:http://dx.doi.org/10.1037/0012-1649.31.4.697

Chan, D. W. (2010). Developing the impossible figures task to assess visual-spatial talents among chinese students: A rasch measurement model analysis. *Gifted Child Quarterly, 54*(1), 59-71.

Cherney, I. D., & Collaer, M. L. (2005). Sex differences in line judgment: Relation to mathematics preparation and strategy use. *Perceptual and Motor Skills, 100*(3), 615-627. doi:http://dx.doi.org/10.2466/PMS.100.3.615-627

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology, 7*(1), 20-43. doi:http://dx.doi.org/10.1016/0010-0285(75)90003-1

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. *Visual information processing* (pp. xiv, 555). Oxford, England: Academic, Oxford.

Cooper, L. A., & Shepard, R. N. (1975). Mental transformation in the identification of left and right hands. *Journal of Experimental Psychology: Human Perception and Performance, 1*(1), 48-56. doi:http://dx.doi.org/10.1037/0096-1523.1.1.48

Dean, A. L., & Harvey, W. O. (1979). An information-processing analysis of a piagetian imagery task. *Developmental Psychology, 15*(4), 474-475. doi:http://dx.doi.org/10.1037/0012-1649.15.4.474

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*(10), 850-855. doi:http://dx.doi.org/10.1111/j.1467-9280.2007.01990.x

Frick, A., Daum, M. M., Walser, S., & Mast, F. W. (2009). Motor processes in children's mental rotation. *Journal of Cognition and Development, 10*(1-2), 18-40. doi:http://dx.doi.org/10.1080/15248370902966719

Frick, A., Daum, M. M., Wilson, M., & Wilkening, F. (2009). Effects of action on children's and adults' mental imagery. *Journal of Experimental Child Psychology, 104*(1), 34-51. doi:http://dx.doi.org/10.1016/j.jecp.2009.01.003

Frick, A., Ferrara, K., & Newcombe, N. S. (2013). Using a touch screen paradigm to assess the development of mental rotation between 3 1/2 and 5 1/2 years of age. *Cognitive Processing, 14*(2), 117-127. doi:http://dx.doi.org/10.1007/s10339-012-0534-0

Funk, M., Brugger, P., & Wilkening, F. (2005). Motor processes in children's imagery: The case of mental rotation of hands. *Developmental Science, 8*(5), 402-408. doi:http://dx.doi.org/10.1111/j.1467-7687.2005.00428.x

Geiser, C., Lehmann, W., & Eid, M. (2006). Separating "rotators" from "nonrotators" in the mental rotations test: A multigroup latent class analysis. *Multivariate Behavioral Research, 41*(3), 261-293. doi:http://dx.doi.org/10.1207/s15327906mbr4103_2

Guilford, J. P., Fruchter, B., & Zimmerman, W. S. (1952). Factor analysis of the army air forces sheppard field battery of experimental aptitude tests. *Psychometrika, 17*, 45-68. doi:http://dx.doi.org/10.1007/BF02288795

Hamm, J. P., Johnson, B. W., & Corballis, M. C. (2004). One good turn deserves another: An event-related brain potential study of rotated mirror-normal letter discriminations. *Neuropsychologia, 42*(6), 810-820. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2003.11.009

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: Similarities and differences between mental rotation and mental folding. *Cognitive Processing, 14*(2), 105-115. doi:http://dx.doi.org/10.1007/s10339-013-0544-6

Harris, J., Newcombe, N. S., & Hirsh-Pasek, K. (2013). A new twist on studying the development of dynamic spatial transformations: Mental paper folding in young children. *Mind, Brain, and Education, 7*(1), 49-55. doi:http://dx.doi.org/10.1111/mbe.12007

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139.

Jolicoeur, P. (1988). Mental rotation and the identification of disoriented objects. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, *42*(4), 461-478. doi:http://dx.doi.org/10.1037/h0084200

Kail, R. (1986). The impact of extended practice on rate of mental rotation. *Journal of Experimental Child Psychology*, *42*(3), 378-391. doi:http://dx.doi.org/10.1016/0022-0965(86)90032-9

Kail, R., Carter, P., & Pellegrino, J. W. (1979). The locus of sex differences in spatial ability. *Perception & Psychophysics*, *26*(3), 182-186. doi:http://dx.doi.org/10.3758/BF03199867

Kail, R., Pellegrino, J., & Carter, P. (1980). Developmental changes in mental rotation. *Journal of Experimental Child Psychology*, *29*(1), 102-116. doi:http://dx.doi.org/10.1016/0022-0965(80)90094-6

Khooshabeh, P., Hegarty, M., & Shipley, T. F. (2013). Individual differences in mental rotation: Piecemeal versus holistic processing. *Experimental Psychology*, *60*(3), 164-171. doi:http://dx.doi.org/10.1027/1618-3169/a000184

Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., & Alpert, N. M. (1998). Mental rotation of objects versus hands: Neural mechanisms revealed by positron emission tomography. *Psychophysiology*, *35*(2), 151-161. doi:http://dx.doi.org/10.1017/S0048577298001516

Kyllonen, P. C., Lohman, D. F., & Snow, R. E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology*, *76*(1), 130-145. doi:http://dx.doi.org/10.1037/0022-0663.76.1.130

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*(4), 940-949. doi:http://dx.doi.org/10.1037/0012-1649.35.4.940

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6), 1479-1498. doi:http://dx.doi.org/10.2307/1130467

Lizarraga, M. L. S. d. A., & Ganuza, J. M. G. (2003). Improvement of mental rotation in girls and boys. *Sex Roles, 49*(5-6), 277-286. doi:http://dx.doi.org/10.1023/A:1024656408294

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature.* (Tech. Rep. No. 8). Stanford, CA: Aptitude Research Project, School of Education, Stanford University.

Marmor, G. S. (1975). Development of kinetic images: When does the child first represent movement in mental images? *Cognitive Psychology, 7*(4), 548-559. doi:http://dx.doi.org/10.1016/0010-0285(75)90022-5

Marmor, G. S. (1977). Mental rotation and number conservation: Are they related? *Developmental Psychology, 13*(4), 320-325. doi:http://dx.doi.org/10.1037/0012-1649.13.4.320

Martin, N. A. (2006). *Test of visual perceptual skills* (3rd ed.). Novato, CA: American Therapy Publications.

Milivojevic, B., Johnson, B. W., Hamm, J. P., & Corballis, M. C. (2003). Non-identical neural mechanisms for two types of mental transformation: Event related potentials during mental rotation and mental paper folding. *Neuropsychologia, 41*(10), 1345-1356. doi:http://dx.doi.org/10.1016/S0028-3932(03)00060-5

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin, 27*, 272-277.

Moore, D. S., & Johnson, S. P. (2008). Mental rotation in human infants: A sex difference. *Psychological Science, 19*(11), 1063-1066. doi:http://dx.doi.org/10.1111/j.1467-9280.2008.02200.x

Mullis, I. V. S., & Martin, M. O. (.). (2013). *TIMSS 2015 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Nevala, J., & Lyytinen, H. (2000). *Sanaketjutesti [test of word chains]*. Jyväskylä, Finland: Niilo Mäki Institute, Jyväskylä University.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Parsons, L. M. (1987a). Imagined spatial transformation of one's body. *Journal of Experimental Psychology: General, 116*(2), 172-191. doi:http://dx.doi.org/10.1037/0096-3445.116.2.172

Parsons, L. M. (1987b). Imagined spatial transformations of one's hands and feet. *Cognitive Psychology, 19*(2), 178-241. doi:http://dx.doi.org/10.1016/0010-0285(87)90011-9

Piaget, J., & Inhelder, B. (1971). *Mental imagery in the child: A study of the development of imaginal representation* (P. A. Chilton Trans.). London: Routledge & K. Paul.

Quinn, P. C., & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science, 19*(11), 1067-1070. doi:http://dx.doi.org/10.1111/j.1467-9280.2008.02201.x

Räsänen, P. (2000). *NMART—Arithmetic reasoning test*. Jyväskylä, Finland: Niilo Mäki Institute. Unpublished material.

Räsänen, P., & Aunola, K. (2007). *The 3-minutes basic arithmetic test*. Jyväskylä, Finland: Niilo Mäki Institute. Unpublished.

Rochat, P., & Hespos, S. J. (1996). Tracking and anticipation of invisible spatial transformation by 4- to 8-month-old infants. *Cognitive Development, 11*(1), 3-17. doi:http://dx.doi.org/10.1016/S0885-2014(96)90025-8

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119*(1), 5.

Searle, J. A., & Hamm, J. P. (2012). Individual differences in the mixture ratio of rotation and nonrotation trials during rotated mirror/normal letter discriminations. *Memory & Cognition, 40*(4), 594-613. doi:http://dx.doi.org/10.3758/s13421-011-0172-2

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology, 93*(3), 604-614. doi:http://dx.doi.org/10.1037/0022-0663.93.3.604

Shepard, R. N., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology, 3*(2), 228-243. doi:http://dx.doi.org/10.1016/0010-0285(72)90005-9

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701-703. doi:http://dx.doi.org/10.1126/science.171.3972.701

Suokas, M. (2009). *Lukusujuvuustestin kehittäminen ja arviointi.* Psykologian pro gradu. University of Jyväskylä.

Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology, 22*(7), 996-1013. doi:http://dx.doi.org/10.1002/acp.1420

Thurstone, L. L. (1938). *Primary mental abilities* University of Chicago Press: Chicago.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352-402. doi:http://dx.doi.org/10.1037/a0028446

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250-270. doi:http://dx.doi.org/10.1037/0033-2909.117.2.250

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817-835. doi:http://dx.doi.org/10.1037/a0016127

Wechsler, D. (2003). *WISC-IV: Wechsler intelligence scale for children* NCS Pearson, Incorporated.

Woodcock, R. W. (1997). The woodcock-johnson tests of cognitive Ability—Revised.

Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review*, *15*(4), 763-771. doi:http://dx.doi.org/10.3758/PBR.15.4.763