THE THRESHOLD HYPOTHESIS APPLIED TO SPATIAL SKILL AND MATHEMATICS

By

Daniel Freer

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Educational Psychology and Educational Technology —Doctor of Philosophy

2017

ABSTRACT

THE THRESHOLD HYPOTHESIS APPLIED TO SPATIAL SKILLS AND MATHEMATICS

By

Daniel Freer

This cross-sectional study assessed the relation between spatial skills and mathematics in 854 participants across kindergarten, third grade, and sixth grade. Specifically, the study probed for a threshold for spatial skills when performing mathematics, above which spatial scores and mathematics scores would be significantly less related. This study explored the relation of spatial and mathematics skills in three ways: age, type of spatial assessment, and whether the mathematics was new or already taught. Students were assessed on a battery of spatial and mathematical assessments. Spatial thresholds were discovered using a segmented regression analysis based on age and based on whether the mathematics was new or already taught, but not based on type of spatial assessment. Specifically, a spatial threshold was found in mathematics overall, in third and sixth grade students, in new and already taught mathematics for kindergarteners, in new mathematics for third graders, and in new mathematics for sixth graders. Thus, a spatial threshold was found in these instances: the relation between spatial and mathematics scores significantly decreased for students who had spatial scores above a threshold, relative to students below the spatial threshold.

TABLE OF CONTENTS

LIST OF TABLES……………………………………………………………………...v

LIST OF FIGURES……………………………..…………………………………………vi

KEY TO SYMBOLS………………………………………………………………………vii

Introduction 1

Overview and Basic Aims 3

Literature Review..6
Relation of Spatial Skill and Mathematics...6
Spatial Skill and Mathematics are Correlated..6
Students with Mathematical Deficits have Poor Spatial Skill 9
Spatial Skill Predicts Mathematical Performance 10
Underlying Processes that May Connect with Spatial Skill 11
Visual Spatial Training 13
Visual Spatial Training: Effects on Mathematics 15
Threshold Hypothesis 16

Hypothesis 23
Aim 1: General Threshold for Spatial Skill 23
Aim 2: Aspects of Spatial Skills 24
Aim 3: Threshold for New versus Already Learned Mathematics 26
Aim 4: Age Dependent Spatial Thresholds 27

Method 29
Measures 30
Mental Rotation 30
Visual Spatial Working Memory...31
Test of Visual Motor Integration 31
Block Design 32
Map Reading..32
Perspective Taking...33
Place Value. 33
Word Problems. 34
Calculation...34
Missing Term Problems/Algebra...35
Number Line Estimation 35
Fractions. 36
Supplemental Sixth Grade Tests..37

Statistical Analysis..38

Standardization....39
Finding a Threshold....40
Other Potential Analyses....41

Results....43
Aim 1: General Threshold for Spatial Skill Statistical Analysis....43
Aim 2: Aspects of Spatial Skills Statistical Analysis....44
Aim 3: Threshold for New versus Already Learned Mathematics Statistical Analysis....45
Kindergarten....46
Third Grade....47
Sixth Grade....48
Aim 4: Age Dependent Spatial Thresholds Statistical Analysis....49
Spline Analysis....50

Discussion....53
Evidence for the Threshold Hypothesis....53
Evidence Against the Threshold Hypothesis....55
Aim 1: General Threshold for Spatial Skill....55
Aim 2: Aspects of Spatial Skills....58
Aim 3: Threshold for New versus Already Learned Mathematics....59
Kindergarten....61
Third Grade....62
Sixth Grade....64
Aim 4: Age Dependent Spatial Thresholds....66

Conclusions....69
Bidirectionality of Spatial and Mathematics Skills....71

Implications....73

Limitations....75

APPENDIX....76

REFERENCES....98

LIST OF TABLES

Table 1: Mathematics and Spatial Assessments Based on Grade Level……….……………….… 77

Table 2: Correlations between Mathematics and Spatial Scores………………………………. 78

Table 3: Means and Standard Deviations of Low and High Spatial Groups…..……………..… 79

Table 4: Threshold Hypothesis Results…………………………………………………...…… 82

Table 5: Mathematics Problems Split into New and Already Taught…………….…………... 83

Table 6: Comparison of Segmented Regression and Spline Analysis Fit……………………… 84

LIST OF FIGURES

Figure 1: Mental rotation task examples……….………………………………………………… 85

Figure 2: Map reading task……..……………………………………………………………… 86

Figure 3: Kindergarten perspective taking task example……….……………………………… 87

Figure 4: Third and sixth grade perspective taking task example……………………………… 88

Figure 5: Overall Spatial Threshold for Overall Mathematics Model Graphs………….……… 89

Figure 6: Overall Spatial Threshold for Overall Mathematics…………….…………………… 91

Figure 7: Block Design Threshold for Overall Mathematics…………………………………… 91

Figure 8: Perspective Taking Threshold for Overall Mathematics……..……………………… 92

Figure 9: Mental Rotation Threshold for Overall Mathematics………..……………………… 92

Figure 10: Spatial Threshold for New Mathematics in Kindergarteners……….……………… 93

Figure 11: Spatial Threshold for Old Mathematics in Kindergarteners………..……………… 93

Figure 12: Spatial Threshold for New Mathematics in Third Graders………….……………… 94

Figure 13: Spatial Threshold for Old Mathematics in Third Graders…………..……………… 94

Figure 14: Spatial Threshold for New Mathematics in Sixth Graders………….……………… 95

Figure 15: Spatial Threshold for Old Mathematics in Sixth Graders…...……………………… 95

Figure 16: Spatial Threshold for Mathematics in Kindergarteners…….……………………… 96

Figure 17: Spatial Threshold for Mathematics in Third Graders………….…………………… 96

Figure 18: Spatial Threshold for Mathematics in Sixth Graders………….…………………… 97

KEY TO SYMBOLS

VSWM Visual Spatial Working Memory

EFA Exploratory Factor Analysis

CFA Confirmatory Factor Analysis

IQ Intelligence Quotient

VMI Visual Motor Integration

Introduction

It is commonly assumed that more is better when it comes to skills, ability, and human performance. This is the idea behind most educational efforts, wherein students receive training on the assumption that it will lead to higher levels of skill or better performance. It is also often assumed that if children are trained in school, the training will transfer to other tasks. Indeed, if two skills are related, it makes sense to expect training in one domain to lead to improvement in both. However, this idea has not always worked out. For example, some studies show that spatial training on one task does not always lead to either improvement on other similar spatial tasks or transfer to mathematics tasks, even though mathematics ability and spatial ability are correlated (Uttal et al., 2013). Such lack of transfer after training has also been shown for working memory and executive function training. Because training and transfer skills relate to each other, it makes sense that training on one skill would improve the other. When this does not occur, it begs the question, why not?

One reason may be that once a person reaches a certain skill level, additional training will not lead to further improvement in related areas. This idea is similar to the idea that an upgrade of a computer's hardware will not significantly improve performance of software that already runs on that computer: for example, a calculator and supercomputer can both calculate 2 x 2 equally well despite the supercomputer's superior hardware. In other words, there may be thresholds for some skills that, once passed, do not lead to further improvement in correlated skills.

The idea that simultaneous improvement in correlated skills can stop improving simultaneously is called the Threshold Hypothesis, and has been studied in many areas including IQ and creativity, IQ and success, ability level and career choice, lifestyle and vocational interests, working memory and language learning, and mathematical ability and education level (Kim, 2005; Lubinski, Webb, Morelock, & Benbow, 2001; Runco, & Albert, 1986; Wai, Lubinski, & Benbow, 2009). For example, researchers have established that IQ is positively correlated with creativity up to a score of 120. However, with IQ scores above 120, the positive correlation no longer holds; past this threshold more intelligence does not positively correlate with correspondingly higher levels of creativity. In the present study, I will look for evidence of the existence of a threshold for spatial skill and mathematics. This relation is important to investigate because the results of spatial training studies have been mixed, particularly in terms of transfer to mathematics. The Threshold Hypothesis could offer an important framework for understanding these discrepant findings.

Overview and Basic Aims

The current study is a secondary analysis of the data reported in Mix, Levine, Cheng, Young, et al. (2016). This study assessed students' spatial and mathematical skill in an exploratory factor analysis. Spatial skills are used to remember, understand, and manipulate the relation between objects in space. In a follow-up study with a second wave of data collection, a confirmatory factor analysis was conducted and yielded the same pattern of results (Mix, Levine, Cheng, Young et al., 2016). The two studies found that certain aspects of spatial skill relate to certain aspects of mathematical skill and that these relations varied with age. In this dissertation, I performed a series of analyses on data from both studies to determine whether a threshold exists in any of these relations. I used segmented regression analysis in which participants were divided into high and low spatial skill groups to determine if there is a threshold at which the slope between spatial and mathematics skill significantly changed, and at what level of spatial skill the threshold exists. This approach has been used previously to find a threshold in the relation of IQ and creativity (Jauk, Benedek, Dunst, & Neubauer; 2013). I investigated the potential of a spatial threshold in four specific ways.

A. To determine whether overall visual spatial skill and mathematics are positively correlated only up to a certain threshold, I examined the overall visual spatial scores concurrently with the overall mathematics scores. After combining the visual spatial and mathematics data into composite scores, I performed a segmented regression analysis that splits participants into high visual spatial and low visual spatial groups. I hypothesized that there will be a threshold at which spatial and mathematical skills are no longer correlated.

B. To see whether the same thresholds apply to different tasks, I examined three specific visual spatial skills separately: perspective taking, mental rotation, and block design. I examined assessments of these three skills to determine whether they have similar or different thresholds relative to each other and relative to visual spatial skills as a whole. I performed a segmented regression analysis that splits participants into high visual spatial and low visual spatial groups for each task. I hypothesized that each different spatial skill will have a different threshold that relates to overall mathematics scores. The spatial scores assessed in this analysis may have different thresholds, though some may exhibit no threshold if the participants' scores remain at the same slope and level of correlation.

C. To see whether the same thresholds apply to already known and new mathematics, I examined mathematics by splitting it up into mathematics that was already taught and mathematics that was not taught. This will be referred to as already learned and new mathematics. The content of new and already learned mathematics was determined by the common core curriculum and whether or not teachers had covered the material being assessed. I performed a segmented regression analysis to compare the already learned material to the unlearned material to see if there is a different threshold for visual spatial skill. I hypothesized that new mathematics will have a spatial threshold, whereas already learned mathematics will not.

D. To compare different age groups to determine if there are different thresholds of visual spatial skill that are dependent on age, I assessed participants from three grade levels: kindergarten, third grade, and sixth grade. I sought to determine (1) whether there is a threshold in each age group, (2) the strength of the threshold, and (3) if each age group has different thresholds. This was also done by standardizing the students' scores, averaging them together to create composite scores based on grade level for spatial and mathematics scores, then finding the threshold score using a segmented regression analysis. I hypothesized that younger students will have a higher spatial threshold to be skilled at mathematics than older students, based on findings in prior research that has shown that younger people use spatial skills more when solving mathematics problems (Holmes & Adams, 2006, Holmes et al., 2008, McKenzie et al., 2003, Rasmussen & Bisanz, 2005, Hecht et al., 2001, Noël et al., 2004 and Passolunghi et al., 2007).

Literature Review

Relation of Spatial Skill and Mathematics

In this first section, a review of the literature will demonstrate a relation between spatial skill and mathematics. This relation is well established, though as we will see, the precise mechanism of this relation is still being explored. There are several sources of evidence for this relation. Mathematics scores and spatial scores are correlated, and spatial scores have predictive influence on mathematics scores (Alloway & Passolunghi, 2011; Gathercole & Pickering, 2000; Johnson, 1998; Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki, 2003; Lachance & Mazzocco, 2006; Markey, 2010; Mazzocco & Myers, 2003; Mix et. al., 2016). Another source of evidence for this relation is the shared cognitive processes between spatial tasks and mathematics tasks that are presumed by prior research, which may be the reason for the correlation and prediction relations (Tartre, 1990; Booth & Thomas, 1999; Van Garderen & Montague, 2003). Third is a relation between training in spatial skill and its effects on mathematics scores, which has had mixed results (Cheng and Mix, 2014; Hawes, Moss, Caswell, & Poliszczuk, 2015). The mixed results on spatial training's effect on mathematics and explaining why that may be are the focus of this paper.

Spatial Skill and Mathematics are Correlated

Research has indicated that visual spatial skill correlates with mathematical skill; this has been verified with several different measures of visual spatial skill and mathematics (Gathercole & Pickering, 2000; Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki, 2003; Mix et. al., 2016). This relation has been demonstrated in several different age groups with varying types of visual spatial assessments and in varying types of mathematics (Alloway & Passolunghi, 2011; Johnson, 1998; Lachance & Mazzocco, 2006; Markey, 2010; Mazzocco & Myers, 2003). The research showing correlation between spatial skill and mathematics is just one way spatial skill and mathematics are related, and this relation is still being investigated by researchers.

A study by Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki (2003) found that children with a high level of visual spatial skill perform better on counting tasks. In this study, preschool children were given assessments to measure visual spatial working memory, which contained mental rotation tasks, static visual spatial working memory tasks, and dynamic visual spatial working memory tasks. An early numeracy test was given to the participants as well, in which the children completed eight sections that examined comparison, classification, correspondence and seriation, use of numerals, structured counting, resultative counting, and general knowledge of numbers. The researchers found that counting skills are correlated with visual spatial ability and concluded that they are related. Since these results were found in preschoolers, this relation is not a result of education, which is significant because it indicates that the relation between visual spatial ability and mathematics skill might be innate rather than taught in school. The authors concluded that children need physical objects to develop counting skills and that the physical act of counting reflects the use of visual spatial skills and trains the children.

Mix et al. (2016) performed an exploratory factor analysis on students' spatial skills and mathematical skill. This study was done with kindergarteners, third, and sixth graders using the same data set used in this paper. Exploring several different aspects of spatial skills and mathematics problems, researchers found that there are shared sub-skills among spatial skills and mathematics. For kindergarteners, the spatial skills of mental rotation and block design were strongly related to mathematics; for sixth graders, the spatial skills of visual spatial working memory (VSWM) and visual motor integration (VMI) were strongly related to mathematics. However, for third graders there were no significant cross-domain loadings in their analysis.

This study provokes one of my research questions, which is related to whether the relation between spatial ability and mathematical skill manifests in different ways at different grade levels.

Alloway & Passolunghi (2011) found similar results and showed that seven- and eight-year old students with high visual spatial skill had overall better mathematical performance than those with low visual spatial skill. While this study examined working memory as a whole, it found that visual spatial skills and verbal skills correlated with mathematics scores in seven-year-olds and only visual spatial ability correlated with mathematics scores in eight-year-olds. The measures of visual spatial working memory were three: an odd-one-out task where the children view and identify which of three shapes is the odd one out, then remember that shape's location; the Mr. X task in which children determine if two different figures are holding a ball in the same hand by looking at two pictures; and a spatial recall task where the participants are shown two shapes and try to determine whether the two shapes are the same. To assess mathematical skills, the participants were given Italian AC-MT tests that included number operations, quantity discrimination, number production, and number ranking. Arithmetic abilities were assessed using the numerical operations subtest of the Wechsler Objective Numerical Dimensions test. The correlation between visual spatial working memory and mathematical skills was found in number operations, number production, number ranking, and numerical operations. The researchers concluded that visual spatial skill is used as a mental blackboard for students to solve mathematics problems. Gathercole & Pickering (2000), Meyer, Salimppor, We, Geary, & Menon (2010), and Raghubar, Barnes, & Hecht (2010) also found that visual spatial skill positively correlates with better mathematics performance. Other research has also shown that specific visual spatial tasks correlate with general mathematics scores from kindergarten to 12th grade (Johnson, 1998; Lachance & Mazzocco, 2006; Markey, 2010; Mazzocco & Myers, 2003).

The research discussed has demonstrated a correlational relationship between visual spatial skills across many tasks and mathematical skills and across many different ages; but the established relationship does not end there: researchers have also shown that students with mathematical and arithmetic deficits also have deficits in certain measures of visual spatial skills.

Students with Mathematical Deficits have Poor Spatial Skill

Not only does having a high level of visual spatial skill correlate with students who have high mathematics scores, there is also a correlation with students who have arithmetic deficits and low visual spatial scores. The deficits in mathematics are shown using curriculum assessments, which are important in the academic achievements of students. Six- to seven-year-olds who failed to achieve expected levels on the national curriculum assessments in mathematics also scored low on visual spatial assessments, specifically in static and dynamic matrices tests as well as in static maze assessments (Gathercole & Pickering, 2000).

Fourth and fifth graders who had both reading and mathematics disabilities scored significantly lower than children with just reading disabilities in visual spatial tests, specifically on static matrices assessments (van der Sluis, van der Leij, & de Jong, 2005). Even when general intelligence was accounted for, the children with only the reading disability performed significantly better on the visual spatial task than the children with reading and arithmetic disabilities (van der Sluis, van der Leij, & de Jong, 2005). The authors suggest that the deficit in visual spatial skill is associated with the arithmetic disability, although they do not posit a causal direction of this association.

McLean & Hitch (1999) looked at nine-year-olds with arithmetic and mathematical difficulties to determine if they were also having difficulties in visual spatial skills. The students took the Graded Arithmetics-Mathematics Test, which identified students who were poor in arithmetic. The students in the poor arithmetic group were then matched with an age-matched control and an ability-matched control of students who were a grade younger. Visual spatial skill was measured by using a visual matrix span and a Corsi blocks span; the poor arithmetic group scored significantly worse on Corsi blocks than the age-matched group did. The authors concluded that the deficits in visual spatial skills are important factors in the lower arithmetic scores. This leads to the question, because deficits in visual spatial skills are important factors in arithmetic, do the deficits in visual spatial skills predict mathematical competence?

Spatial Skill Predicts Mathematical Performance

Researchers have demonstrated that the relationship between visual spatial skills and mathematics goes beyond correlation, showing further that visual spatial scores predict future mathematical achievement in students. This has been shown across several different ages and in many different aspects of mathematics. A prediction of general mathematics scores has been demonstrated in studies by Geary (2011) that examined the predictive power first graders achievement to see how it influences growth in mathematics in fifth grade. The researchers found that visual spatial scores, specifically a span task block recall, predicted mathematics achievement.

Beyond predicting future achievement, spatial skill's also predicts mathematical skills across different age groups in several types of visual spatial skills and mathematical knowledge. De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière (2009) demonstrated that visual spatial skills predict mathematics performance in first graders; measures of block recall and visual patterns predicted first graders' mathematics scores 4 months later. This study used two measures of visual spatial skills: block recall, which is associated with the spatial subsystem of visual spatial ability, and the visual patterns test, which is related to the visual subsystem. The spatial subsystem measure was the better predictor of mathematics performance later on which, consistent with previous studies, establishes a relationship between the spatial subsystem and mathematics (Hegarty & Kozhevnikov, 1999; Garderen & Montague, 2003).

Bull, Espy, & Wiebe (2008) examined how first graders' visual spatial skills predicted mathematics in third grade. Visual spatial short-term memory span, as measured by Corsi block scores, was the largest predictor of mathematics scores, more so than verbal working memory, inhibition, and planning and monitoring. The authors conclude that there is an "importance of a good understanding of spatial relations and the importance of being able to manipulate visual-spatial material in working memory as critical to mathematical achievement" (Bull, Espy, & Wiebe, 2008, pg. 223), which suggests that students need to have visual spatial skills to learn and perform some mathematical concepts. This leads us to the questions: why do people need visual spatial skills to learn mathematics, and do specific cognitive processes relate to different mathematical problem solving processes?

Underlying Processes that May Connect with Spatial Skill

The fact that visual spatial skill correlates with and predicts mathematical skill suggests that they may have shared cognitive processes. Visual spatial skill is the ability to imagine and manipulate figures – both mentally and in physical space. This ability can manifest in many different ways, one of which is solving mathematics problems. This aspect of visual spatial skill, which enables the manipulation of figures, requires similar cognitive processes to those needed to solve mathematics problems.

A study by Van Garderen & Montague (2003) demonstrated that students who excel in mathematics create mental representations of the problems they are solving. This was determined by testing sixth grade students who were categorized as gifted, average achievers, or learning disabled. Students that are more gifted would use visual spatial representation of the problem to help them find a solution, where they would have a picture in the mind of the problem they were solving. This spontaneous use of visualization by gifted students and its correspondence with higher achievement on the mathematics task suggest an inherent connection between the two skill sets. When using visual spatial representation, students in all categories were more likely to get the answer correct. This study is significant because it suggests the importance of visual spatial skills on mathematical performance, across cognitive ability level. It also provides evidence of shared cognitive processes between visual-spatial skill and mathematics skill because students who visualized the problems performed better on the mathematics task.

Booth & Thomas (1999) found similar verbal results when they gave students the same problems in one of three formats: verbally, with a picture, or with a diagram. The participants were split up into high visual spatial and low visual spatial groups; and even though both groups scored the same on standardized mathematics tests, the high visual spatial group performed significantly better on problems with a diagram. The authors concluded that students with higher visual spatial skill are better at some types of mathematics problems that require a higher level of visualization including oral, picture, and diagram problems.

Tartre (1990) concluded that "spatial orientation skill appears to be used in specific and identifiable ways in the solution of mathematics problems" (pg. 227). The role of spatial skills in this study was explored by splitting tenth grade students into low spatial and high spatial skill groups, and then testing how well students performed on geometric problems and whether they created visual representations of the problem. The results suggest that having visual spatial skills aids in learning and performing mathematics. Specifically, visual spatial skills were helpful in students' accurately estimating figures, adding marks to show relationships, and solving problems when there are not visual frameworks provided.

Visual spatial skill has been shown to correlate with and predict mathematical skill, and it has been shown to be used in solving and learning mathematics, particularly for younger children. If students who possess more visual spatial skill are better at performing and learning mathematics, does training in spatial skill transfer to greater facility in mathematics?

Visual Spatial Training

Visual spatial skill's importance in academic areas such as mathematics has led to the question of whether we can train students' visual spatial skill. Researchers have tried many different training regimens to improve visual spatial skills and so far the results of this research have been mixed. Research has shown that training a specific visual spatial skill results in visual spatial skill improvement. A meta-analysis by Baenninger & Newcombe (1989) found that visual spatial training improved visual spatial skills; specifically, they looked at test-specific visual spatial training, which was meant to train a specific spatial sub-skill, relative to general visual spatial training, which was meant to improve all aspects of spatial skill. Since this study was a meta-analysis, many different types of visual spatial training were included as long as the training met the researchers' minimum duration requirement. The authors found that the test-specific training improved the participants' scores significantly more than the general visual spatial training. Other studies have also found that visual spatial training can improve a specific visual spatial task (Ehrlich, Levine, & Goldin-Meadow, 2006; Heil, Rosler, Link, & Bajric, 1998; Hsi, Linn, & Bell, 1997; Kail, 1986; Newcombe & Frick, 2010; Sorby & Baartmans, 2000; Uttal et al., 2013; Vasta, Knott, & Gaze, 1996). These results prompt the question: does visual spatial training generalize to other non-trained visual spatial tasks?

Research has found mixed results whether visual spatial training will generalize. A study by Terlecki, Newcombe, & Little (2008) showed that participants who were trained using videogames demonstrated learning that transferred to spatial tasks such as the clock task, which shows participants a clock and asks them to match the first clock to a second identical clock that has been rotated out of a group of clocks. This study demonstrated that the effects lasted over several months. The videogame that participants trained on in this study was Tetris and involved either 2-D or 3-D representations of the game depending on the participants' group. To measure improvement in participants’ mental rotations, they were pre-tested and post-tested using the Mental Rotation Task, which had the participants mentally rotate 2-D and 3-D objects. To measure transfer, participants were assessed using the Guilford-Zimmerman Spatial Visualization Task, which had participants change the rotation of a clock, and the Surface Development Test, which had participants match a folded piece of paper with a 2-D rendering of it. The researchers found the effects of playing Tetris on transfer to be significant in both transfer tests and the results were maintained over several months. Other researchers also concluded that visual spatial training can generalize (DeLisi & Cammarano, 1996; Wallace & Hofelich, 1992)

However, some research has failed to replicate the visual spatial training generalization results. A study by Kail & Park (1990) found that mental rotation training did improve 11- and 20-year-old participants’ mental rotation skills; however, the training did not transfer to other areas. Participants repeatedly performed a mental rotation task and improved their time; however, there was no corresponding improvement in a memory search assessment. The conclusion of the authors was that transfer would only occur with stimuli that are similar to the training stimuli. Individual researchers have found different results in terms of visual spatial training. A meta-analysis by Uttal et al. (2013) tried to determine whether visual spatial training is effective, and what kind of transfer visual spatial training can produce. This meta-analysis found evidence for near and medium transfer to trained and related tasks, but did not find significant evidence of far transfer to tasks in other untrained tasks. While research has not come to a consensus, these authors concluded that visual spatial training could be beneficial for mathematics, science, and engineering despite mixed current results.

Visual Spatial Training: Effects on Mathematics

Researchers have begun looking at far transfer effects of visual spatial training in mathematics, and so far, the results have been mixed. A study by Cheng and Mix (2014) found that visual spatial training transfers to mathematics. The researchers trained first graders in a mental rotation task where participants would match two parts of a shape with its whole. Then, they assessed students' mathematical skills by testing number fact problems, calculation problems, and missing term problems. The authors found that the mathematics scores improved in participants who received the visual spatial training, particularly on missing term problems, and concluded that visual spatial working memory capacity increased after training and that this expanded visual spatial working memory capacity supported students' calculation performance. These findings are significant because in addition to providing further evidence of a link between visual spatial training and mathematics performance, they suggest that visual spatial ability is a skill that can be improved with training.

In contrast, Hawes, Moss, Caswell, & Poliszczuk (2015) found that spatial training did not improve mathematics scores in 6- to 8-year olds. This study trained students in mental rotation and found improvement in mental rotation after training, as well as marginally significant improvement in untrained mental rotation tasks. However, researchers found no evidence of transfer to mathematics. Consistent with these findings, a study by Ferrini-Mundy (1987) found that visual spatial training was not effective at improving college students' calculus scores. Researchers have found mixed results on whether or not visual spatial training is effective at transferring to different visual spatial tasks, as well as non-related mathematics tasks. Transfer results were mixed while using a variety of different spatial training and different mathematics assessments across various ages. The discrepancies in the literature could be due not training in the appropriate spatial skill or testing in the appropriate mathematics skills, but the variety already tested has shown no pattern when it came to spatial training or mathematics assessment choice.

What could be the reason for this discrepancy in the literature? One explanation is that there is a threshold beyond which additional improvement in spatial skill does not lead to changes in mathematics. In other words, spatial training might only improve mathematics scores if students lack a minimal level of spatial skill. Students who already possessed the minimal level of spatial skill necessary might not improve their mathematics scores after additional spatial skills training. If so, then a section of the population for each of the published training studies could already have high enough visual spatial skill that further training would not influence mathematics scores.

Threshold Hypothesis

Threshold Theory states that when two skills or abilities are related, one skill has relation to the other up to a certain threshold. After one skill reaches above the threshold, its relation to the other skills either diminishes or disappears.

This type of relationship has been established in the relation of IQ and creativity: IQ correlates to improved creativity up to a certain extent (Runco, & Albert, 1986; Kim, 2005). This relation may also exist between spatial skills and mathematics. Different researchers state that you need different levels of IQ, from 100 to 130, to be creative, but researchers who support this theory all claim that once a person has a high enough IQ, any more would be superfluous when it comes to creativity (Runco, & Albert, 1986; Kim, 2005). This relationship exists because a certain level of IQ is needed to be creative. Given that the cognitive process of visual spatial skill is used in solving mathematics; could there then be a threshold of spatial skills that students would need to be skilled in mathematics?

Research has shown that there is a threshold hypothesis for intelligence and creativity. Fuchs-Beauchamp, Karnes, and Johnson (1993) found that intelligence is needed for creativity in preschoolers. This study measured 496 preschoolers' intelligence with a Slosson Intelligence test or Standford-Binet Intelligence Scale and creativity was measured with a Thinking Creatively in Action and Movement Scale. The authors found that IQ and creativity were significantly positively correlated when students IQ’s were under 120 and not related when their IQ was greater that 120; this provides evidence for the Threshold Hypothesis marking the threshold to be creative at 120 (Fuchs-Beauchamp, Karnes, & Johnson, 1993). While other researchers have also shown that there is a threshold for intelligence to be creative, Jauk, Benedek, Dunst, & Neubauer (2013) empirically tested at what IQ level the intelligence threshold exists. The researchers had 297 participants ranging from 18 to 55 years old. They used four subtests from the Intelligence Structure Battery to measure intelligence and the Inventory of Creative Activities and Achievements to measure creativity. To find where the threshold is in intelligence the researchers used a segmented regression analysis and found that the Threshold Hypothesis applies to intelligence and creativity, and that the threshold is 120.

While some studies have shown promise with applying the Threshold Hypothesis to academic areas, some research has not supported the Threshold Hypothesis. A study that followed 320 of the top .01% of gifted students for 10 years found that there was a positive correlation between being gifted and pursuing a doctoral degree; this correlation remained strong at every level and there was no threshold (Lubinski, Webb, Morelock, & Benbow, 2001). The study by Lubinski, Webb, Morelock, & Benbow, (2001) also looked at gifted individuals' income, publications, patents and tenure and found that there was a strong correlation at high levels of ability, but still no threshold. The Threshold Hypothesis was also analyzed for students who did not have a proclivity for science, technology, engineering, or mathematics careers. When students excelled in verbal skill but lacked mathematical skill, this correlation was found throughout all students with no threshold which would indicate certain career choices (Wai, Lubinski, & Benbow, 2009).

An analysis of several different threshold hypothesis studies by Robertson, Smeets, Lubinski, & Benbow (2010) concluded that there is no threshold for several different areas where researchers have attempted to apply the Threshold Hypothesis. This study looked at the Threshold Hypothesis in gifted mathematics and science graduate students' cognitive abilities, vocational interests, and lifestyle preferences when applied to career choice, performance, and persistence. However, the authors of this study were analyzing only the top 1, 5, or 10 percent, and did not choose the threshold they were analyzing based on any empirical analysis such as a segmented regression analysis. Thus, this study’s provision of conflicting evidence does not definitively disprove the Threshold Hypothesis. The Threshold Hypothesis has been shown to exist in several different areas; therefore, it could exist in other areas as well.

Just as thresholds have been found for other areas of human cognition, it is possible there is a threshold beyond which additional training in spatial skills would have no effect on mathematics performance. The cornerstone of the relation between spatial skills and mathematics is the correlation that has been well documented (Gathercole & Pickering, 2000; Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki, 2003; Mix et. al., 2016); this strong correlation leads researchers to ask other questions about the relation between spatial skill and mathematics. Other questions researchers have asked and found evidence for is if spatial skills predicts mathematics scores (Bull, Espy, & Wiebe, 2008; De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière, 2009; by Geary, 2011). These two relations found between spatial skills and mathematics demonstrate a strong connection between the two, and raise the question what other relations can exist. Researchers further probed this relation by investigating if spatial skills are used in mathematics problems, and found that spatial skill was used when performing mathematics (Booth & Thomas, 1999; Tartre, 1990; Van Garderen & Montague, 2003). However, does that mean spatial training will improve mathematics? Indeed, will unlimited spatial training lead to unlimited gains? Or is there a threshold beyond which additional training in spatial skill has no further effect on mathematics? On the Threshold Hypothesis, if students already have the spatial skills required to perform mathematics, then the relation between spatial skills and mathematics will decrease for that group; while students who do not have the required spatial skill will have a stronger relation between spatial skill and mathematics. This would be evident if the relation between spatial skills and mathematics decreases once students have enough spatial skill to perform mathematics. While this study does not propose to answer the second question about training students, determining whether such a threshold exists is the first step to further exploring the potential benefits of this relation.

Applying the Threshold Hypothesis in the relation between visual spatial skill and mathematics would change our interpretation of the correlational relationship that has already been established. The established positive correlational relationship shows that as people have higher visual spatial skill, they will be better at mathematics. While this general premise will remain the same, the threshold hypothesis will refine this idea. If the threshold theory applies to visual spatial skill and mathematics, there will be a significant positive correlation between visual spatial skill and mathematics to an extent, but once a person has reached a certain level of visual spatial skill, the correlation with mathematics will decrease or possibly even cease. This would indicate that a certain level of visual spatial skills is needed to learn mathematics, but any greater skill, whether through naturally occurring variation or training, would not have as great an effect on a person's mathematical scores.

A Threshold Hypothesis in visual spatial skill and mathematics could have implications for training in visual spatial skill as well. Researchers have decided to train students in visual spatial skills to see if there is an effect on mathematical skills because they have found a correlation between those two skills. If there is a positive correlation between visual spatial skill and mathematics, then training students in visual spatial skill and testing subsequent mathematical ability would help determine if there is causation. If there is, improving visual spatial skill through training would be related to an increase in mathematics scores.

Broadly, training in visual spatial skill has had mixed results. Different studies show varying levels of effectiveness and of transfer to other areas of visual spatial skill, while some studies show no improvement after training. While visual spatial training has had mixed results, research has looked to see if visual spatial training improves mathematics scores, based on the relation between visual spatial skill and mathematics. If there is a threshold relationship between visual spatial skill and mathematics, then training in visual spatial skill will not be effective in students whose visual spatial skill is above threshold, or if their deficits may be in other areas such as knowledge of the material or attention deficit hyperactive disorder (ADHD). These results would indicate that students should only be trained in visual spatial skill if they are below the needed threshold to learn and perform mathematical tasks.

Further, there are many different types of mathematics and visual spatial skills; therefore, there might be different thresholds for different types of mathematics as they relate to different types of visual spatial skill. These differences would need to be taken into consideration during training so as to train students in the correct visual spatial skill deficits to get to the proper threshold. Thus, more nuanced evidence for the Threshold Hypothesis in this research area would alter our understanding of visual spatial skill and mathematics and change our interpretation of this previously established relation. Not only would such evidence for the Threshold Hypothesis change our understanding, it would also change interventions that have been designed to improve students' visual spatial skill.

Mathematics and spatial skill's relation has been established through correlation and prediction, while there are still questions about causation. The correlation and prediction between spatial skills and mathematics exists through different aged participants and different tasks. However, spatial skills are more correlated with mathematics scores at younger ages, and different spatial skills have different levels of correlation to different types of mathematics. This study aims to address the questions about causal factors between spatial skills and mathematics, specifically why spatial training has not consistently improved mathematics scores. The established relations between mathematics and spatial skills could exist for many different reasons; one explanation is shared cognitive processes. While exploring how spatial skills are used to solve mathematics, Van Garderen & Montague (2003) and Tartre (1990) found evidence of shared cognitive processes of both skills. The use of spatial skills to solve mathematics problems led me to see spatial skills as being similar to the hardware of a computer, while mathematics skills are analogous to the software. If hardware is sufficient to run a given software, any upgrades to the hardware will not significantly improve that software. However, on this analogy, you do need a certain level of hardware to run certain software. If you have the spatial skills to perform mathematics, improving your spatial skills will not significantly influence your mathematical performance; however, you do need a certain level of spatial skill to perform mathematics. The importance of understanding the thresholds of spatial skills when it comes to learning mathematics is that it both improves our understanding of the relation between spatial skills and mathematics, and explains why different spatial training studies have had mixed results.

Hypothesis

Is there a level of visual spatial skill that people need to be skilled at mathematics, and is having more than that unnecessary for learning and performing mathematics? To find out, I carried out a series of segmented regressions that probe for thresholds at several levels of analysis, from general to specific. Below, I describe the rationale for each analysis and the hypothesis it will address. I explain the general method for the study, including sampling and statistical approach; then I explain the specific analysis for each aim.

Aim 1: General Threshold for Spatial Skill

The first question I intend to answer is whether there is a general threshold for spatial skill, after which an increase in spatial skill does not positively correlate with better mathematics scores. Mix et al. (2016) found no evidence that spatial skills were subdivided into components such as spatial visualization, or dynamic versus static spatial skills. This may be because spatial skill is a general skill; and, therefore, spatial skills as a whole may have a threshold when it comes to mathematics.

To test this I created a general spatial score for students. Several components of spatial skills were assessed in this data collection and were used to determine if general spatial skills have a threshold when performing mathematics. General spatial skill was tested because we know that spatial skills are correlated to each other. To create a general spatial score I created a Z-score for all spatial assessments and a composite general mathematics score combining all mathematics assessments for every student. These two scores were then compared using a segmented regression analysis.

The standardization of IQ scores has aided the investigation of Threshold Hypothesis by helping all researchers agree on and understand the different scores. For example, different researchers believe people need different levels of IQ to be creative, with some stating that level is 130 while others posit it at only 100. Although they disagree about the threshold, there is a common understanding of the way IQ is scored. This is not the case for visual spatial skill. Because there is no standardized score for visual spatial skill, my results may be more difficult to interpret, but the use of Z-scores will address this shortcoming to some extent. The Z-score will allow for the comparison of assessment scores that have different types of problems and different distributions of scores.

Aim 2: Aspects of Spatial Skills

What makes the relationship between visual spatial skill and mathematics more complicated are the many different ways to measure and represent visual spatial skills, because "spatial ability is not a monolithic and static trait, but made up of numerous sub-skills, which are interrelated among each other and develop throughout your life" (*What is spatial ability?*, Johns Hopkins University, pg. 1). The relationship with mathematics becomes more complicated as both the type of mathematics children learn and the way they solve problems change over time; these changes may require students to use different cognitive processes and strategies to solve mathematical problems.

Visual spatial skill is divided into different sub-skills and different researchers have created different divisions. For example, a study by Hegarty & Kozhevnikov (1999) split visual spatial skill into two categories: schematic representation, which codes the spatial relationship between objects; and pictorial representations, which codes the appearance of objects. The schematic representations correlated with mathematical problem solving, while pictorial representations did not. Van Garderen & Montague (2003), using sixth grade students, split their participants into the same two categories of visual spatial skill, and found similar results in the relationship between schematic representations and mathematical performance.

Linn & Peterson (1985) have split up visual spatial skill further into three different categories: mental rotation, spatial perception, and spatial visualization. Mental rotation involves rotating either two- or three-dimensional figures. Spatial perception involves a person determining the orientation of an object relative to them. Spatial visualization requires a person to analyze the relationship between different spatial representations. While each of these categories fits into visual spatial skill, all have different roles and relationships with mathematics.

Researchers have demonstrated that visual spatial skill has a relationship with mathematics; however, each of the different sub-skills of visual spatial ability has a different relationship with mathematics (Hegarty & Kozhevnikov, 1999; Linn & Peterson, 1985; Van Garderen & Montague, 2003). There are many different visual spatial assessments that measure the different sub-skills of visual spatial ability, and many have shown to correlate with or predict mathematics aptitude or performance. This intricate relationship between visual spatial skills and mathematics results in a non-uniform threshold for visual spatial skill and mathematics with thresholds at different levels. For example, students learning a new mathematical concept will rely more heavily on visual spatial skills than an adult who already knows said mathematical concept and has alternative procedural knowledge of how to solve certain mathematical problems (Hubber, Gilmore, & Cragg, 2014).

The current study's data were collected by giving students six different assessments of visual spatial skill. The different sub-skills measured by each assessment will have differing strengths in their relation with mathematics (Hegarty & Kozhevnikov, 1999; Linn & Peterson, 1985; Van Garderen & Montague, 2003). For example, the schematic sub-skill of visual spatial ability, which encodes spatial relationships, shows a stronger relationship with mathematics than the pictorial sub-skill, which encodes visual appearance (Hegarty & Kozhevnikov, 1999). With a variety of mathematics assessments at different grade levels, the type of mathematics will also affect the threshold of visual spatial skill needed to be skilled at mathematics. Varieties of visual spatial assessments were used to determine which, if any, visual spatial sub-skills indicate a threshold requirement to be skilled at mathematics.

Any relationship between mathematics and visual spatial skill is dependent on the specific skill or assessment being examined. Therefore it is important not only to look at how visual spatial skill, as a whole, relates to general mathematical skill, but it is also important to look at specific visual spatial assessments and compare them to different types of mathematics. I hypothesize that different spatial sub-skills will have different thresholds for mathematics, and some sub-skills may have no threshold.

Aim 3: Threshold for New versus Already Learned Mathematics

If a student has already learned certain content, there may be either minimal or no visual spatial threshold required to be able to perform those mathematical problems. The reason for this could be that once students learn a mathematical concept and they learn the procedure to solve it as a mathematical problem, they will no longer need to rely on visual spatial skills (Hubber, Gilmore, & Cragg, 2014). Hubber, Gilmore, & Cragg (2014) demonstrated this by testing adults where the participants would solve mathematics problems while doing visual spatial tasks, and found that adults do not rely as much as children on visual spatial skills while performing mathematics, which was shown by their difference in strategy choices. This study took adults and tested them on a dynamic spatial n-back task, which uses the visuospatial sketchpad and central executive cognitive capacities, to specifically assess adults counting, decomposition, and direct retrieval strategies. The researchers found that adults rely on spatial ability to perform mathematics, but only because they use domain general executive skills. These results suggest that adults do not use spatial skills as much as younger people do when solving mathematics problems and instead rely on more general skills. Their explanation for this is that adults have developed mathematics strategies, while children do not have strategies, are still developing their strategies, or have less efficient strategies.

These findings differed from prior research, which had found that performing mathematics relies on visual spatial skills (Dumontheil & Klingberg, 2012; Heathcote, 1994; Reuhkala, 2001; Simmons et al., 2012; Trbovich & LeFevre, 2003). Hubber, Gilmore, & Cragg's (2014) study used adults, while the other studies used children; they suggested that adults rely less on visual spatial skills because they have developed different strategies to perform mathematics that do not rely on visual spatial skills. To test if students rely on visual spatial skills on problems for which they have not developed comprehensive learning strategies, which is procedural knowledge that students use to perform mathematics and rely less on spatial skills,, the problems in the different age groups were split up into new content and content already learned by that age group. I suggest that the new content problems will have a higher visual spatial threshold needed to perform the mathematics.

Aim 4: Age Dependent Spatial Thresholds

It is possible that the threshold for spatial skill and mathematics shifts with development. Research has shown that younger children's mathematics scores correlate more with visual spatial skills than older students do. Research by De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière (2009) demonstrated that first graders' mathematics scores correlate more with visual spatial skills than second graders by specifically assessing students engaged in both a block recall task, which has participants remember a sequence of blocks, and a visual patterns test, which has participants remember what parts of a grid are filled up. Both tasks correlated with number knowledge, understanding of operations, simple arithmetic, word problems, and measurement. The authors concluded that middle school students use physical objects more to perform mathematics problems, and that using physical objects to solve those problems requires visual spatial working memory. Other research has also found this trend where younger children's visual spatial scores correlate more with mathematics than do older students' (Holmes & Adams, 2006, Holmes et al., 2008, McKenzie et al., 2003, Rasmussen & Bisanz, 2005, Hecht et al., 2001, Noël et al., 2004 and Passolunghi et al., 2007)

I hypothesize that to be skilled at mathematics; the threshold for visual spatial skill is higher for younger children than for older children and adults. The reasoning for this is that younger children rely more on visual spatial skill in mathematics than do older children (De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière, 2009). Students at several different age groups were assessed in their visual spatial skills to determine if younger students require a higher threshold of visual spatial skill to be skilled at mathematics than older students.

Method

This study is a secondary analysis of a dataset that includes 854 children from 33 schools in the Midwest region of the United States. Each student completed a battery of up to 15 tests (depending on age) that measured mathematics and spatial skills (see Table 1). A total of 952 had parents who gave consent out of 3,749 parents that were contacted from the 33 schools. Of the children who consented, 98 children were excluded because (1) tests were missing due to absences, they declined to participate, or their schools withdrew participation after consents were received (*n* = 26); (2) data that were corrupted due to experimenter error (*n* = 68); or (3) from children who were part of a non-typical population (English language learners, special education students, etc.) (*n* = 4). The sample of 854 children with usable data was split into three age groups: kindergarten (*n* = 275, 131 boys, mean age = 6.04, *SD* = .40), third grade (*n* = 291, 142 boys, mean age = 9.04 years, *SD* = .41) and sixth grade (*n* = 288, 131 boys, mean age =11.74 years, *SD* = .44). Over the course of two weeks, the participants were tested in three or four 1-hour sessions each depending on age. Some tests were administered individually (details provided below under "Measures"), while other tests were administered in small groups (*n* = 4-6), or to entire classrooms for sixth grade participants (*n* = 25-30). Dividers were used in kindergarten and third grade group tests to prevent participants from copying. The test order was blocked and counterbalanced by individual versus group administration. Further, tests were administered in blocks, and the order of tests in the block was random. Participants in all three grades received a decorated folder as a reward for participation.

Measures

The procedures and materials for the specific measures are described below, as well as the reliabilities for each grade. Some measures were standardized tests and have published reliabilities that I report here. For the others, I computed reliabilities from our own data using Cronbach's alpha.

Mental Rotation

Mental rotation (adapted from Neuberger, Jansen, Heil, & Quaiser-Pohl, 2011 and Peters, Laeng, Latham, Jackson, et al., 1995). Two types of mental rotation were based on Vandenberg and Kuse’s (1978) mental rotation task. For kindergarten and third grade participants, small groups of participants were shown four two-dimensional figures of capital letters and asked to point out which two letters were the same as the target letter. The two items that match could be rotated to overlap and match the target, whereas the other two targets did not match because they were mirror images. This assessment was introduced by giving the participants four practice items on a laptop screen. The children received feedback on whether or not they were correct when answering. The participants were shown animations that demonstrated the correct answerers being rotated to match the target letter. After the practice problems, children completed 16 test items in a paper booklet (kindergarten $\alpha = .72$; third grade $\alpha = .87$). The version for the sixth grade participants was the same, except that participants were shown 12 items consisting of perspective line drawings of three-dimensional block constructs. Two of the drawings shown could be rotated to match the picture to the target ($\alpha = .79$). Children received credit for answering each item correctly only if both matches were identified.

Visual Spatial Working Memory

Visual spatial working memory (adapted from Kaufman & Kaufman, 1983). For this assessment, for each trial participants were shown a 14 cm x 21.5 cm grid that was divided into squares (e.g., 3 x 3, 4 x 3, etc.), with objects displayed in random positions in the grid. The items were made more difficult by making the grid bigger (up to 5 x 5) and adding objects (up to nine). For every trial the grid with objects was displayed for five seconds, and then the objects were removed with only the grid remaining. The participants would then indicate by marking an 'X' on their own paper grid where they thought the objects were on the grid. The grids for the stimulus and participants' response sheet were marked with lines. The grid and objects were shown on a laptop computer and children marked their responses in a paper test booklet. The assessment was introduced to participants using three practice problems for which the participants received feedback on these initial practice problems by comparing their answers to the stimulus display. The test trials (n = 19 for kindergarten, n = 15 for third grade, n = 29 for sixth grade) began immediately after the second practice trial. The reliabilities were based on the current data, because the initial publication was several decades ago (α = .74, .63, and .82 for kindergarten, third grade, and sixth grade, respectively).

Test of Visual Motor Integration

Test of visual motor integration (VMI, 6th ed., Beery & Beery, 2010). Participants copy a line drawing of a geometric shape onto a blank sheet of paper on each trial, for example a 3-D cube. Participants answer 18-24 trials depending on the participants' age, where the figures being drawn would become more complex. This test was administered in small groups. The reliability of the VMI, based on a split-half correlation (reported in the test manual), was .93.

Block Design

Block design (WISC-IV) (Wechsler, Kaplan, Fein, Kramer, et al., 2004). Participants were shown a printed square figure comprised of red and white sections, and they re-created that figure using small cubes with red, white, and half red half white sides. This test was administered individually after instructions were read from the WISC-IV manual. It became increasingly more difficult, and children completed more or fewer problems depending on their basal and ceiling performance. The reliability coefficient reported in the WISC-IV manual for the Block Design subtest is between .83 and .87 depending on age group.

Map Reading

Map reading (adapted from Liben & Downs, 1989). Participants in kindergarten and third grade completed 14 test trials in which they were shown a full color three-dimensional model town with buildings, roads, a river, and trees. The model was 10 by 10 inches, with the tallest structure being 0.50 inches tall. The task for sixth grade participants was similar, but the locations were shown using a full-color aerial photograph of an actual town (for eight trials total). For each trial, the location in the photograph was shown to the participants as a location on the aerial photograph. Participants would then identify where on the map the location was in the photograph. Difficulty of the questions was changed by varying the scale ration (1:1, 1:2.5) and degree of rotation between the photograph or model, and the map (0 to 180). The problems were ordered from easiest to most difficult. Participants received feedback on the first three questions to ensure that the participants understood the task. Younger participants were tested individually, but the sixth grade participants completed the assessment in groups. The reliability of this task was kindergarten α =.56; third grade α =.72; sixth grade α =.57.

Perspective Taking

Perspective taking (Frick, Mohring & Newcombe, 2014; Hegarty & Waller, 2004; Kozhevnikov & Hegarty, 2001). Participants completed one of two perspective taking assessments that required children to imagine a scene from different perspectives. The kindergarten and third grade participants' version (adapted from Frick et al., 2014) had the participants view a set of Playmobil figures in specific arrangements. The participants were shown four pictures and asked which picture was taken from each figure's perspective. The questions' difficulty changed depending on the number of objects in the picture and the angles of view. The participants completed 4 practice problems and were given feedback to ensure they understood the problem; the test had 27 questions in addition to the 4 practice questions. The sixth grade participants' version (adapted from Kozhevnikov & Hegarty, 2001) was similar. Participants were shown six to eight objects arranged in a circle. The participants were then asked to envision themselves next to one of the objects while facing another object in the circle, then draw an arrow to a third object that was in the circle to indicate their angle of view from that perspective. Sixth grade participants completed two practice problems with feedback, followed by 12 test items. The reliability of this test was kindergarten $\alpha = .56$; third grade $\alpha = .87$; sixth grade $\alpha = .84$.

Place Value.

Participants in kindergarten and third grade completed an assessment of 20 questions that require children to compare, order, and interpret multi-digit numerals that was different than the sixth grade participants' assessment (e.g., "Which number is in the one's place", as well as match multi-digit numerals to their expanded notation equivalents (342 = 300 + 40 + 2). Reliability on this experimenter-constructed measure was $\alpha = .79$ at kindergarten and $\alpha = .79$ at third grade. The sixth grade participants completed similar concepts on the Rational Numbers subtest (CMAT) (α =.94 reported for 12-year-olds in test manual). The sixth grade participants were also asked to compare, order, and interpret written numbers, but these items consisted of a mix of multi-digit numbers, fractions, and decimals.

Word Problems.

The kindergartener and third grade participants were tested using 12 word problems from the Test of Early Mathematics Ability-Third Edition (TEMA-3, Ginsburg & Baroody, 2003) (kindergarten α =.70; third grade α =.63). The TEMA-3 is a test of numerical skills, such as cardinality, calculation, and commutativity. This assessment was administered individually to the participants following the instruction set out in the TEMA manual. Although the participants completed the entire TEMA-3, only the word problems were analyzed here. To measure performance on word problems among sixth grade students, we used the Problem Solving subtest from the CMAT (α =.89 reported for 12-year-olds in the test manual).

Calculation.

Participants received a group test consisting of 12-28 age-appropriate arithmetic problems (kindergarten: n = 16, α = .76; third grade: n = 12, α = .69; sixth grade: n = 28, α = .77). For kindergarteners the problems were one- to four-digit whole number addition and subtraction problems. Third grade participants had one- to four-digit whole number addition and subtraction problems, as well as four whole number multiplication and division problems with one to three digits. The sixth graders' calculation assessment had 28 problems that had all four operations. Sixteen of the problems used whole numbers up to five digits and 12 of the problems used decimals.

Missing Term Problems/Algebra.

Participants' performance on missing term problems was a part of a separate measure, because prior research suggested a causal relation between spatial skill and performance on this type of mathematics problem (Cheng & Mix, 2014). In a missing term problem the participant finds the solution to a calculation problem where a problem is shown and one of the terms is missing (e.g., $X + 5 = 17$). Kindergarten and third grade participants completed missing term problems, but sixth grade students did not complete them because they would not be challenging (n = 8 items, kindergarten: $\alpha = .61$; third grade: $\alpha = .71$). Sixth grade students used the Algebra subtest from Comprehensive Mathematics Ability Test (CMAT, Pro-Ed, 2003) to measure a similar mathematical skill (although algebra has additional components and cognitive demands). The CMAT is standardized for the age range 7 to 19 years of age and was administered in groups (n = 10-25). The reliability for the CMAT Algebra subtest was reported in the test manual as $\alpha = .88$ for 12-year-olds.

Number Line Estimation

Number line estimation (Booth & Siegler, 2006; Siegler & Opfer, 2003). Participants were tested in groups of four to six. They were initially shown a line with a numeral at each end (e.g., 0 and 100). After participants were shown a card with a written numeral on it, the participants then marked on the number line where the number should go. The task was introduced by the experimenter asking children to mark where a "small" and "big" number would go, to ensure that participants understood that smaller numbers went left and larger numbers went right. However, experimenters did not provide feedback about the correct position of numerals on the number line. The specific numerals shown to participants varied by age group. Specifically, kindergarteners placed the numerals 4, 17, 33, 48, 57, 72 and 96 on a 0-

to-100 number line (even-odd reliability: $r = 0.37$); third grade students placed 3, 103, 158, 240, 297, 346, 391, 907 on a 0-to-1000 number line (even-odd reliability: $r = 0.32$); and sixth grade students placed 25,000, 61,000, 49,000, 5,000, 11,000, 2,000, 15,000, 73,000, 8,000, 94,000 on a 0-to-100,000 number line (even-odd reliability: $r = 0.56$). While these reliabilities were relatively low, related work shows linear R^2 values for subsets of number line estimates vary widely (see Young and Opfer, 2011). In addition, when we computed the reliabilities using error rate instead of linearity, they were well above conventionally accepted levels (kindergarten $\alpha = .74$, third grade: $\alpha = .87$, sixth grade: $\alpha = .86$). In our analyses, we focused on linearity because this variable captures internally consistent placements (i.e., sets or responses that were linear relative to each other even if they were not mapped onto the number line itself) that might be missed if absolute distance to the target were used (i.e., if all the responses were skewed to the high or low end of the number line but were nonetheless, increasing linearly); however, we also investigated whether task relations changed when error rates were used instead (see footnote 2).

Fractions.

Fraction concepts are typically introduced in third grade and become a major part of the mathematics curriculum by sixth grade (e.g., *Common Core State Standards for Mathematics*). Therefore, fraction items were not included in the kindergarten assessments. For third grade assessments four items were included that tested fraction equivalence and simple calculation with common denominators ($\alpha = .56$). Two assessments were used to estimate sixth grade participants' understanding of fractions. One assessment had 22 items that tested comparison, calculation with and without common denominators, and calculation with mixed numbers ($\alpha = .75$). The second assessment was a version of number line estimation task in which the number line was anchored with 0 and 1, and the quantities to be placed were all fractions (i.e., 1/4, 1/19, 2/3, 7/9, 1/7, 3/8, 5/6, 4/7, 12/13, 1/2) (split half reliability for linearity: r = .40, for error rate =.78) (e.g., Fazio, Bailey, Thompson, & Siegler, 2014). The two fraction measures for sixth grader participants were entered separately to better evaluate the relations involving number line estimation.

Supplemental Sixth Grade Tests.

The breadth of mathematics skills increase significantly during middle school and Mix et al. (2016) thought it possible that skills not measured in younger children might be related to spatial skill in sixth grade. Therefore they assessed sixth grade participants' performance on two subtests from the CMAT - Charts and Graphs (α =.91) and Geometry (α =.77) - to tap those additional skills. In the Charts and Graphs subtest, data in graphic form are shown to students. Then they are asked questions involving interpretation of the graphic information (e.g., when shown a bar graph with the number of packages mailed each day for a week, students are asked how many packages were mailed on Wednesday, and whether more packages were mailed on Friday or Monday). For the Geometry subtest participants were asked to identify geometric forms (angles, lines, solids, etc.), solve for unknown angles, calculate perimeter, area, and volume, and so forth.

Statistical Analysis

The basic aims of this study were to evaluate whether spatial skill and mathematics are more strongly related among low versus high performing students. Such a pattern would be evident if the correlation between spatial skill and mathematics changed slope at a certain level of spatial performance, at what could be called a spatial performance "threshold." To make these evaluations, I used three main analyses. The primary analysis for identifying thresholds is to compare slopes for one segment to the other, I used segmented regression analyses and, in particular, the Davies Test. The Davies Test measures significant changes in slope. If there is a significant change in slope, this would support the hypothesis that the Threshold Hypothesis exists in this data set. I also evaluated whether spatial skill and mathematics were significantly correlated in each of the two segments using Pearson's correlations. As I will outline below, there are several possible outcome patterns that may or may not indicate a threshold. By examining both changes in slope and the degree of correlation in each segment, I can accurately determine which pattern has been obtained. Last, I evaluated whether the correlations in these two segments differed in strength using a Fisher's Z-score. This provided very specific information about the nature of the relation between spatial skill and mathematics for each segment. As stated above, the Davies test is the measure that determines if the Threshold Hypothesis exists in this data set, the additional Fisher's Z-score analysis can provide evidence for a stronger threshold. For example a significant change in slope alone, as measured by the Davies Test would be a weak threshold; while, a significant change in slope in addition to a significant change in correlation, as measured by the Fisher's Z-score would indicate a strong threshold. The additional analysis enables the categorization of results into no threshold, weak threshold, and strong threshold.

Hypothetically, there are multiple possible patterns for segmented regression analyses. Three of these possibilities are depicted in Figure 5. Figure 5a shows a significant correlation with no threshold. Figure 5b shows a threshold wherein the low performing group has a significantly higher slope between spatial skill and mathematics but the high performing group does not. Figure 5c shows the reverse: a threshold wherein the high performing group has a significantly higher slope between spatial skill and mathematics but the low performing group does not. Of course, there are many other variations possible: for example, both groups could have significant correlations that may or may not differ in slope, or there could be differences in slope but either non-significant or significant correlations for each segment, or there could be no difference in slope but significant differences in the strengths of the correlations for each segment. Although these nuances were explored and included in my interpretations, the main overall pattern that would be most consistent with my hypothesis is depicted in Figure 5b.

Standardization

Unlike IQ, there is no standardized measure for visual spatial skill that every researcher uses. Therefore, a visual spatial composite score was created that incorporated all of the visual spatial assessments that I was evaluating for each analysis. Before calculating the composite scores, each individual score was transformed into a Z-score. Z-scores were used because they permitted comparisons across measures that may have had a different numbers of questions or different score distributions. Once each student's scores were standardized, the visual spatial scores and mathematics scores were grouped together, with the assessments in each group being dependent on the research question, to test each research aim. For example, to answer the research question about whether thresholds varied for different age groups, the composite scores included all visual spatial and mathematics assessments given to children in a specific grade level.

Finding a Threshold

Historically, in research on IQ and creativity, investigators chose an arbitrary cut-off to use as the threshold (Guilford, 1967). For example, early work took a standardized IQ score of 120 to be the threshold for creativity (Jauk, Benedek, Dunst, & Neubauer, 2013); however, this cut-off ultimately proved controversial because there was no empirical basis for it (Jauk, Benedek, Dunst, & Neubauer, 2013). Similarly, there is no agreed upon threshold for spatial skill. However, it is possible, and perhaps better, to locate the threshold empirically. The 'segmented' package in R statistical software provides a mechanism for finding thresholds in regression data (Mueggo, 2008). This package uses an iterative search to test several possible breakpoints and select the one with the largest change in slope to find the ideal breakpoint.

To evaluate the threshold and determine whether there is a significant change in slope, in the segmented package of 'R' the Davies Test and segmented function was used. The Davies Test evaluates the hypothesis that there is a threshold (i.e., a significant change in slope) and the segmented function measures where the ideal breakpoint is. The Davies test takes several different potential threshold values and tests whether or not there is a significant change in the correlation slope at each value. The estimated breakpoint, signified by Ψ, is a good measure to determine potential thresholds because it vanishes when there is no change in slope, this can only occur when the correlation coefficient is equal to 1 or -1. The segmented function then takes the various values for Ψ found in the iterative search and tests which segment has the most significant change in slope. Ψ values are compared a null Ψ value that is assigned as the mean spatial score. The Ψ value that has the lowest p-value is labeled as the threshold as described in this paper. This value is considered the ideal breakpoint, after this statistical analysis another analysis is run that tests if the breakpoint is a significant change in slope between the two segments. The Davies test then measures for significant change in slope for a value inside of the confidence interval of the ideal breakpoint.

The Davies test does not automatically test the ideal breakpoint; instead it samples the data set a certain number of times. The number of times the data set is sampled is selected by the researcher. I initially started with 10 samples, but changed it if the Davies breakpoint was not within the ideal breakpoint's confidence interval. Changing the number of samples can change the sampling points, with the goal of measuring for a significant change in slope within the confidence interval of the ideal breakpoint. To determine whether each segment represented a significant correlation between spatial skill and mathematics a Pearson's correlation test was used. Each segment was analyzed to determine if the correlation was significant. To evaluate whether the two segments' correlations were significantly different from each other a Fisher's Z-score was used. In each analysis the data was split up into two different segments based on the location of the threshold as determined by Ψ, this analysis measures if there was a significant change in correlation between the two segments.

Other Potential Analyses

Several other statistical analyses might have been possible to include, but were not because they did not address the research questions as well as segmented regression analysis. A cluster analysis will split participants in different groups based on their spatial and mathematics scores. However, having participants in more than two groups does not properly address the original hypothesis. The original hypothesis is that students need a certain level of spatial skill to be skilled at mathematics; students either have enough spatial skill or they do not. Thus, to conduct a cluster analysis would provide interesting information about the potential existence of multiple groups with differing levels of spatial and mathematical ability, but it would not provide insight about the potential existence of a threshold.

A curvilinear relationship could also be examined, but that would not fit the hypothesis either. A curvilinear relationship is still correlational for participants who go beyond the spatial threshold. The theory states that any excess spatial skill would not impact mathematics scores, while the curvilinear relationship would determine whether the scores started getting worse. What this study aimed to test was whether the positive correlational relationship significantly decreases or disappears after certain spatial scores are reached. A curvilinear analysis would determine if the relationship goes from positively correlated to negatively correlated, while instead we would like to test if it goes from strongly positively correlated, to little or no correlation that can be either positive or negative. Thus, a test for a curvilinear relationship would have little interpretive value when considered in light of the aims of the current study.

Results

Aim 1: General Threshold for Spatial Skill Statistical Analysis

I first asked whether there was a general threshold, irrespective of grade level or task type. For this analysis aggregate measures were used, in which assessments were grouped across grade levels into two broad categories: one category that included every spatial assessment and one category that included every mathematics assessment. Each assessment was standardized using a Z-score and grouped together to find the mean. Next, I carried out a segmented regression analysis to determine whether and at what level a threshold for visual spatial skill relative to mathematics skill might exist. The segments and correlations are summarized in Tables 2 and 3.

A threshold was detected at 1.23 standard deviations below the mean. The slope between mathematics and spatial scores for students below this point was 1.81 (n = 34) and above this point, the slope was .59 (n =1557). According to the Davies test, this change in slope was significant (p = .01; CI = -1.45 - -1.10); however, both the low and high performing groups had significant correlations between spatial skill and mathematics, and a Fisher's Z-score comparing the two correlations was not significant (z = -1.09, p = .138). Taken together, these results indicate there was a threshold, but the relation between spatial skill and mathematics was not correlationally different for one segment or the other. Thus, my hypothesis was supported in part, but not as strongly as it could have been if the two segments had had different correlations, with the lower spatial skill group having a stronger correlation (see Figure 6 and Table 4).

Aim 2: Aspects of Spatial Skills Statistical Analysis

The second research question asked whether there were thresholds for specific spatial assessments when performing mathematics. For this analysis, I chose three spatial tests based on Linn and Peterson's (1985) categorization of spatial skills: the block design, perspective taking, and mental rotation. Linn and Peterson (1985) categorized spatial skills as mental rotation, spatial perception, and spatial visualization. Block design assessment is meant to represent spatial visualization, perspective taking assessment is meant to represent spatial perception, and mental rotation assessment is meant to represent mental rotation. My hypothesis is that the specific spatial assessments will have significant thresholds where spatial skills and mathematics are related in the lower spatial group, with a significant decrease in slope for the higher spatial group. In addition, the different spatial assessments may have different thresholds. For each assessment, I standardized the test scores using Z-scores and then regressed each one on a composite standardized mathematics score, containing all the mathematics assessments. Next I analyzed the three specific spatial assessments using a segmented regression analysis to determine if there were thresholds for each specific spatial assessment in overall mathematics. There were no significant thresholds, in terms of change in slope, according to the Davies test; however, there were significant changes in correlation for each group according to the Fisher's Z-score. These results did not provide evidence in support of this study's hypothesis

The block design assessment was analyzed for a threshold when looking at all mathematics scores combined into a composite score. A threshold was detected 1.18 standard deviations below the mean. The slope between mathematics and spatial skills for students below this point was .56 (n = 191) and above this point, the slope was .44 (n = 1,401). According to a Davies test, this change in slope was not significant (p = .44; CI = -1.506- -.854); however, a

Fisher's Z-score comparing the two correlations was significant (z = 2.07, p = .039) (see Figure 7 and Table 4).

The threshold for perspective taking in a composite mathematics score indicated a threshold at 2.33 standard deviations below the mean. The slope between mathematics and spatial skills below this point was -1.0 (n = 7) and above this point was .34 (n = 1,585). According to a Davies test, this change in slope was not significant (p = .39; CI = -2.89 - -1.77); however, a Fisher's Z-score comparing the two correlations was significant (z = -8.3, p = .0) (see Figure 8 and Table 4).

The mental rotation assessment was analyzed for a threshold in overall mathematics scores. A threshold was found at .659 standard deviations above the mean. The slope between mathematics and spatial scores for students below this point was .274 (n = 1206) and above this point, the slope was .399 (n = 386). According to a Davies test, this change in slope was not significant (p = .229; CI = -.387 – 1.707); however, a Fisher's Z-score comparing the two correlations was significant (z = -2.41, p = .016) (see Figure 9 and Table 4).

Aim 3: Threshold for New versus Already Learned Mathematics Statistical Analysis

The third research question asked whether there were different thresholds for the relation between spatial skill and mathematics that students have already been taught or have not learned. For this analysis, mathematics problems were split into "new" or "already taught" mathematics based on the common core curriculum (see Table 5). Composite standardized Z-scores were created for the "new" and "already taught" mathematics in each grade. A composite standardized spatial score was tested for a change in slope threshold for both "new" and "already taught" mathematics using a segmented regression analysis. Previous research has shown that spatial skills relate to new and already taught mathematics differently. Specifically that new mathematics uses more spatial skills than already taught mathematics. This suggests that if there is a significant threshold for new or already taught mathematics, the threshold may be different.

This study's hypothesis was that a significant threshold would exist for new mathematics, where mathematics would improve as spatial skills improved, then once spatial scores reached the threshold mathematics scores would no longer improve. For already taught mathematics I hypothesized that there would be little to no correlation between spatial skills and mathematics. I examined this hypothesis separately for each grade level.

Kindergarten

For mathematics problems that were new to kindergarteners a threshold was detected at 1.193 standard deviations above the mean (see Figure 10). The slope between mathematics and spatial scores for students below this point was 1.681 (n = 515) and above this point, the slope was .435 (n = 11). According to a Davies test, this difference in slope was significant (p = .003; 95% CI = .827-1.56 points). The lower spatial skill group had a significant correlation between spatial skill and mathematics while the higher group did not, and a Fisher's Z-score comparing the two correlations was significant as well (z = 2.07, p = .019). Taken together, these results both indicate the presence of a threshold that reflects potentially greater spatial processing for children at lower levels of skill related to new mathematics content (see Figure 10 and Table 4).

For mathematics problems that the kindergarten students were already taught, a threshold detected was at 1.023 standard deviations above the mean. The slope between mathematics and spatial scores for students below this point was 1.047 (n = 506) and above this point, the slope was -2.575 (n = 20). According to a Davies test, this was significant (p < .001; 95% CI = .921-1.124); the lower spatial skill group had a significant correlation between spatial skill and mathematics while the higher skill group did not; and a Fisher's Z-score comparing the two correlations was significant as well (z= 4.56, p = 0). Taken together, these results both indicate a significant threshold for spatial skill such that spatial and mathematics scores are significantly positively correlated up to the threshold; after the threshold spatial and mathematics scores slope significantly decreases and is no longer significantly correlated (see Figure 11 and Table 4).

Third Grade

For mathematics problems that were new to third grade students, a threshold was detected at .191 standard deviations below the mean. The slope between mathematics and spatial scores for students below this point was 1.493 (n = 203) and above this point, the slope was 1.143 (n = 334). According to a Davies test, this was significant (p = .009; 95% *CI* = -.464-.083); both high and low performing groups had a significant correlation between spatial skill and mathematics, and a Fisher's Z-score comparing the two correlations was significant as well (z = -1.91, p = .028). Taken together, these results indicate there was a threshold, where the lower spatial skill group had a higher slope than the higher spatial group; however, the higher skill spatial group had a stronger correlation than the lower group. Thus, my hypothesis was supported in part, but not as strongly as it could have been if the Fisher's z-score had not shown lower skill segment had a stronger correlation than the higher group (see Figure 12 and Table 4).

For mathematics that the third grade students had already been taught, a threshold was detected at .193 standard deviations below the mean. The slope between mathematics and spatial scores for students below this point was 1.045 (n = 151) and above this point, the slope was 1.651 (n = 386). According to a Davies test, this was significant (p < .001; 95% *CI* = -.328 - -.058); however, both low and high performing spatial skill groups had significant correlations between spatial skill and mathematics. A Fisher's Z-score comparing the two correlations was not significant (z = -.04, p = .484). While threshold significant threshold was found, the threshold detected did not support this study's hypothesis, because the lower spatial skills group had a lower slope than the high spatial skills group (see Figure 13 and Table 4).

Sixth Grade

For mathematics problems that were new to sixth grade a student, a threshold was detected at .149 standard deviations below the mean (see figure 14). The slope between mathematics and spatial scores for students below this point was 1.778 (n = 197) and above this point, the slope was 1.318 (n = 332). According to a Davies test, this was significant (p < .001; 95% CI = -.328, -.058); both high and low performing groups had a significant correlation between spatial skill and mathematics, and a Fisher's Z-score comparing the two correlations was significant as well (z = -3.32, p = .0005). Taken together, these results indicate there was a threshold, where the lower spatial skill group had a higher slope than the higher spatial group; however, the higher skill spatial group had a stronger correlation was stronger than the lower. Thus, my hypothesis was supported in part, but not as strongly as it could have been if the lower segment had a stronger correlation than the higher group (see Figure 14 and Table 4.

For mathematics that the sixth grade students had been already taught, a threshold was detected at .278 standard deviations above the mean (see Figure 15). The slope between mathematics and spatial scores for students below this point was .983 (n = 362) and above this point, the slope was 4.685 (n = 167). According to a Davies test, this was significant (p < .001; 95% CI = .239-.317); both low and high performing groups had significant correlations between spatial skill and mathematics; however, a Fisher's Z-score comparing the two correlations was not significant (z = 1.02, p = .154). While these results do indicate there was a threshold, the threshold detected did not support this study's hypothesis, due to the higher spatial skills group having a higher slope. (see Figure 15 and Table 4).

Aim 4: Age Dependent Spatial Thresholds Statistical Analysis

My fourth research aim asked whether there were a different thresholds based on children's age (i.e., grade level). This question was asked because past researchers found the relation between spatial skill and mathematics changes depending on age, specifically that younger students rely on spatial skills more than older students. These differences in spatial skills relation to mathematics depending on age, would suggest that if there was a threshold, that it may change depending on age. Specifically I hypothesize that younger students would have a higher spatial threshold than the older students. For this analysis, composite scores for spatial and mathematics skills were created for each grade level. The composite scores included every spatial or mathematics assessment the students took in each grade level.

For kindergarteners a threshold was detected .303 standard deviations above the mean (see Figure 16). The slope between mathematics and spatial scores for students below this point was .49 (n = 386) and above this point, the slope was .85 (n = 139). According to a Davies test, this was not significant (p = .12; CI = -.173 - .953); both high and low performing groups had a significant correlation between spatial skill and mathematics, and however, a Fisher's Z-score comparing the two correlations was significant (z = -2.54, p = .0055). These results show no indication of a significant threshold. Thus, my hypothesis was not supported (see Figure 16 and Table 4).

For third grade students a threshold was detected 1.29 standard deviations below the mean (see Figure 17). The slope between mathematics and spatial scores for students below this point was 2.282 (n = 23) and above this point, the slope was .581 (n = 514). According to a Davies test, this was significant (p = .008; CI = -1.508 - -1.079); however, both the low and high spatial skill performing groups had significant correlations between spatial skill and mathematics, and a Fisher's Z-score comparing the two correlations was not significant (z = .86, p = .195). Taken together, these results indicate there was a threshold, but the relation between spatial skill and mathematics was not qualitatively different for one segment or the other. Thus, my hypothesis was supported in part, but not as strongly as it could have been if the two segments had had different correlations, with the lower skilled group having a stronger correlation (see Figure 17 and Table 4).

For sixth grade students a threshold was detected .02 standard deviations above the mean (see Figure 18, currently known as Figure 8). The slope between mathematics and spatial scores for students below this point was .763 (n = 287) and above this point, the slope was .425 (n = 242). According to a Davies test, this was significant (p = .04; CI = -.409 - .657); however, both the low and high spatial skill performing groups had significant correlations between spatial skill and mathematics, and a Fisher's Z-score comparing the two correlations was not significant (z = 1.22, p = .111). Taken together, these results indicate there was a threshold, but the relation between spatial skill and mathematics was not qualitatively different for one segment or the other. Thus, my hypothesis was supported in part, but not as strongly as it could have been if the two segments had had different correlations, with the lower skilled group having a stronger correlation (see Figure 18 and Table 4).

Spline Analysis

The present study asked if after a certain spatial threshold, the relation between spatial skills and mathematics decreases; however, there are other ways to measure change in relation of two variables. One could argue that a spline analysis is more appropriate because it can track multiple changes in a given relation, whereas a segmented regression analysis can only track a single change in this relationship. Both the segmented and spline analyses measure changes in slope, which is the key measurement in this studies hypothesis. But, the segmented regression analysis may provide a more direct test of the threshold hypothesis because it indicates whether the relation between spatial skill and mathematics significantly decreases after a certain level of proficiency. In contrast, a spline analysis can reveal multiple changes in the relation between spatial skill and mathematics, but not necessarily whether there is a threshold. Nonetheless, I performed a series of spline analyses to determine whether the pattern of results changes. In every case, the spline analysis was roughly equal to the results revealed by the segmented regressions.

To determine which analysis fit the data better an R^2 was calculated and compared between the two types of analyses. Results were mixed, with the segmented regression analysis being a better fit for more research questions; however, a Steiger's Z-analysis to analyze the difference between fit for the segmented regression and splines analysis found no significant difference between fit for any analysis. Thus, the segmented regression was better research question 1 which analyzed all of the assessments in all of the grade levels (segmented R^2 = 3535; splines R^2 = .3514). For research question 2 which split up the data based on grade level; all grade levels analyses, kindergarteners (segmented R^2 = .3097; splines R^2 = .3061) third graders (segmented R^2 = .3856; splines R^2 = .3852) and sixth graders (segmented R^2 = .3858; splines R^2 = .3850), fit better with a segmented regression analysis. The third research question which split the data up based on spatial assessment had mixed results. Segmented regression analysis fit better for block design (segmented R^2 = .2394; splines R^2 = .2390) and perspective taking (segmented R^2 = .1631; splines R^2 = .1612), while a spline analysis fit better for visual spatial working memory assessments (segmented R^2 = .1964; splines R^2 = .1971). The fourth research question which split the data based on new and already taught mathematics, as well as age, had mixed results. The segmented regression analysis fit better for already taught mathematics for kindergarteners (segmented R^2 = .3754; splines R^2 = .3645) and already taught mathematics for sixth graders (segmented R^2 = .6327; splines R^2 = .6120). A spline analysis fit better for new mathematics for kindergarteners (segmented R^2 = .6876; splines R^2 = .6929), new (segmented R^2 = .8062; splines R^2 = .8084) and already taught (segmented R^2 = .8683; splines R^2 = .8706) mathematics for third graders, and new mathematics for sixth graders (segmented R^2 = .7591; splines R^2 = .7596). In total a segmented regression analysis fit better for eight of the analyses, while a spline analysis fit better for 5 of the analyses; the differences in fit for all analyses were not significant. This shows that both the segmented regression and spline analysis were a good fit and suggests that it was correct to look for the Threshold Hypothesis using a segmented regression analysis due to it being an equivalent fit, and better matches the study's hypothesis.

Discussion

This study was a secondary analysis of data previously used by Mix et. Al (2016). The data collected for the initial analysis was a variety of mathematics and spatial assessments for kindergarten, third, and sixth grade students. Previous research indicated that mathematics and spatial skills are strongly positively correlated and predict one another. Further, it has been noted that spatial skills are used to solve some types of mathematics problems (Tartre, 1990). Because of this relation between spatial and mathematics skills, researchers have investigated whether spatial training would transfer to mathematics skills, but results have been mixed.

I hypothesized that the reason training studies have mixed results is that a positive correlation exists up to a certain threshold, but then the relation is no longer significantly positively correlated. This threshold may exist due to mathematics problems only needing a certain level of spatial skills to solve, beyond which further spatial skills would be excessive. Such a threshold could impact training studies because researchers would not have taken into consideration students who may already have enough spatial skills to perform mathematics. For these students, further training in spatial skills would not impact mathematics scores. This posited threshold may change depending on different factors such as age, spatial skills, and whether the mathematics test questions were new to the students. Therefore, composite scores for both mathematics and spatial skills that represent each of these different factors were compiled to test for different thresholds.

Evidence for the Threshold Hypothesis

Results for the threshold hypothesis were mixed (see Table 4). Evidence of thresholds were found in the overall analysis for all grade levels (see Figure 6), new and already taught mathematics for kindergarteners (see Figures 10 and 11), new mathematics for third grade students (see Figure 12), new mathematics for sixth grade students (see Figure 14), overall third grade students (see Figure 17), and overall sixth grade students (see Figure 18). In each of these cases, the results of the Davies test indicated patterns similar to the idealized version depicted in Figure 5b.

However, the results of analyses that compared the strength of the correlations in each segment (i.e., the Fisher's Z-scores) were not always consistent with the Davies Test results, which measures change in slope. One might predict that whenever there is a large enough change in slope to indicate a threshold; that the correlations would necessarily differ in strength, but this was not always the case (see Table 4). The Davies Test and Fisher's Z-score differed in some cases because of the different methods used in searching for a threshold. The Davies Test uses the change in slope to measure threshold, while the Fisher's Z-score uses strength of correlation to measure threshold. In some cases there can be a significant change in slope between mathematics and spatial scores, but the strength of the correlation does not change despite the change in slope.

The results indicate that only new and already taught mathematics comparisons provided evidence for the threshold hypothesis. For this analysis, the Fisher's Z-score indicated a significant difference in strength of correlation between the two segments, with the lower segment displaying a stronger correlation. For all the others, (i.e., the overall analysis for all grade levels combined, overall for third grade, and overall for sixth grade) the Fisher's Z-scores were not consistent with the Davies Test; there was a significant change in strength of correlation, and both correlations were significant as shown by the Pearson's test. Of those three analyses the lower spatial skills group had a stronger correlation for the third grade analysis (see Figure 17), while the higher spatial skills score group had a stronger correlation for both the overall (see Figure 6) and the sixth grade analyses (see Figure 18). The Davies test outcomes similar to Figure 5b found in these analyses indicates that a breakpoint was found that demonstrates the threshold hypothesis as described in this paper. For these breakpoints the lower spatial skill scoring group had a more positive correlation than the higher spatial skill scoring group.

Evidence Against the Threshold Hypothesis

None of the other analyses provided evidence of thresholds. Already taught mathematics for third grade and sixth grade students as evaluated by the Davies test matched Figure 5c (see Figures 13 and 15); which was a significant threshold, but not a threshold as described by this study's hypothesis. The overall kindergarten scores as evaluated by the Davies test matched Figure 5a (see Figure 16). The results derived from already taught mathematics per the Davies test for third and sixth grade students offered evidence of a significant breakpoint, but not one that typifies the threshold hypothesis as described in the paper: this alternative significant breakpoint suggests that the higher spatial skill scoring group had a more positive correlation than the lower spatial skills group (see Figures 13 and 15). Finally, the results derived from overall kindergarten scores showed no significant change in slope (see Figure 16).

Aim 1: General Threshold for Spatial Skill

The first research question asked if spatial skill is needed to perform all mathematics for all grade levels such that once students reach that point additional spatial skill would not significantly improve mathematics performance. To conduct this analysis a composite score was created that included all grade levels, marking both a composite spatial and a composite mathematics score.

When analyzing all participants' spatial and mathematics assessments in all grades, a spatial threshold was found using the Davies test for change in slope. These results matched Figure 5b, where when students' spatial skills scores reached a certain level, any additional spatial skills would not be as positively associated with mathematics skills scores. In other words, these results suggest that students only need a certain level of spatial skills to be skilled at mathematics, and that once they reach that point further spatial skills do not have as significant an impact on mathematical learning and performance (see Figure 6).

However, the Fisher's Z-score for this overall assessment did not provide additional evidence for the threshold hypothesis. The results of the Fisher's Z-score found the strength of the correlation to be significant in both the high and low spatial skills groups, with no significant difference between correlations (see Table 4).

These results provide qualified evidence for this paper's hypothesis, which states that the threshold hypothesis can be applied to mathematics and spatial skills. One explanation may be that students require a certain level of spatial skills to be skilled at mathematics, and any spatial skill beyond that certain level would not have as big of an impact on mathematics scores. These results conform to the analogy that spatial skills are similar to hardware, and mathematics skills are similar to software; a student (computer) only needs a certain amount spatial skills (hardware) to perform mathematics (software) and any improvement beyond the threshold on spatial skills (hardware) will not improve mathematics (software) performance The students in the higher scoring spatial skills group have exceeded the required level of spatial skills needed to perform mathematics, hence the significant decrease in slope between mathematics and spatial skills.

Another explanation could be that students develop procedural knowledge on how to perform mathematics. Procedural knowledge denotes techniques required to perform mathematics that rely less on cognitive processes such as spatial skills. In this alternative interpretation, the results would still indicate that there is a spatial skills threshold for performing mathematics when looking at kindergarten, third, and sixth grade students taken together. The required spatial skills threshold and a procedural knowledge explanation are not mutually exclusive.

Recall that the main reason to look at these thresholds is to determine if spatial training will be helpful. Previous studies have shown that domain general intelligence correlates with specific aspects of academia, such as the STEM disciplines. Researchers have, in turn, suggested that domain general training would transfer to some of these academic areas. Spatial skills, which have a strong relation to mathematics, have been one specific domain general area that especially stands out. Researchers have suggested that training in spatial skills may transfer to mathematics scores; however their results have been mixed.

The threshold hypothesis could explain why domain general training does not always transfer to specific academic areas, and more specifically why spatial training does not transfer to mathematics. The threshold hypothesis when applied to spatial skills and mathematics states that mathematics and spatial training will have a positive correlation up to a certain threshold, such that once that threshold is exceeded, the slope will decrease significantly. In terms of spatial training, this would mean that spatial training would transfer to mathematics for students whose spatial skill was below the threshold; however, for students whose spatial skills are above the threshold, spatial training would not transfer to mathematics. The results for research from Aim 1 indicate that if you train a student in spatial skills who is already above the spatial skills threshold needed to perform mathematics, then that training would not have as much of an impact as it would have on a lower spatial skills scoring group.

Aim 2: Aspects of Spatial Skills

The second research question asked if there were thresholds for specific spatial skill assessments and mathematics. Previous researchers have split up spatial skills into different categories, with some disagreement as to what those categories are (Hegarty & Kozhevnikov, 1999; Linn & Peterson, 1985; Van Garderen & Montague, 2003). These researchers propose that different categories of spatial skills will have different relations with mathematics (Hegarty & Kozhevnikov, 1999; Linn & Peterson, 1985; Van Garderen & Montague, 2003).

The analyses in this study were done by creating a composite mathematics score and comparing it to the block design, perspective taking, or mental rotation assessments after standardization. These assessments were chosen based off Linn and Peterson's (1985) categorization of spatial skills. When analyzing the data for thresholds, no significant thresholds were found (see Figures 7, 8, 9). These results did not fit with the general threshold found in Aim 1. Due to this discrepancy in results, I chose to perform a segmented regression analysis on every spatial assessment to look for a threshold when applied to overall mathematics scores. None of the specific spatial assessments showed any significant threshold. This leads to the question, why is there a discrepancy in this study's results?

A possible explanation for why specific spatial skill assessments did not show a significant threshold may be that the individual assessments did not ask a sufficient number of questions. This would mean the individual assessments were not sufficiently robust, such that the assessments could not fully tap the construct that is spatial skills or their subdivisions. Assessments should be able to measure the lowest and highest scoring students with enough questions so there is not a significant swing in results if a student misses a single question. To properly accomplish this, assessments need to be lengthy across a range of question difficulty.

While all of the spatial assessments when combined may have had a sufficient robustness to represent the spatial construct, individual assessments may not be adequately robust due to their limited length or a lack of variety in difficulty.

Aim 3: Threshold for New versus Already Learned Mathematics

The third research question asked if there was a different threshold for new mathematics compared to already taught mathematics. This question was asked because prior research has shown that students rely more on spatial skills when performing new mathematics problems than on those they have already learned (Hubber, Gilmore, & Cragg, 2014).

This analysis categorized mathematics problems as new or already taught for each age group. New mathematics comprises mathematics that has not been introduced to students in their grade level, and already taught mathematics comprises mathematics that teachers have already introduced to students in an academic setting. These categorizations conform to the common core curriculum.

Prior research has shown that performing new mathematics relies more on spatial skills than already taught mathematics. Based on this research, I hypothesize that a threshold will exist for new mathematics where spatial and mathematics scores will exhibit a strong and positive correlation below the threshold, while above that threshold spatial and mathematics scores will not be so related. As for already taught mathematics, I hypothesize that there will be no spatial threshold when it comes to already taught mathematics. This result may be due to students utilizing a procedural knowledge when it comes to already taught mathematics.

Three composite scores were created: one for new mathematics, one for already taught mathematics, and one for spatial assessments. A threshold was tested for spatial skill scores when looking at new mathematics and already taught mathematics by looking for significant changes in slope or correlation. Significant thresholds were found, shown by changes in slope, when analyzing new and already taught mathematics scores with kindergarten, third, and sixth grade students. But not all of the significant thresholds support this study's hypothesis (see Figures 10, 11, 12, 13, 14, and 15). Already taught mathematics for third and sixth grade students' Davies test slopes matched Figure 5c; such that while the change in slope was significant, it was not evident of a threshold as described in this paper’s hypothesis (see Figures 13 and 15).

For new mathematics, a spatial threshold was found as predicted by my hypothesis at all grade levels. These results could be due to students only requiring a certain level of spatial skills to perform mathematics. In terms of our previous analogy, if a computer has the hardware to run a given program, upgrading its hardware does not improve that software. For spatial skills, once a student has enough spatial skills to perform mathematics problems, additional spatial skills will not improve performance.

Already taught mathematics for kindergarteners show a similar threshold to the new mathematics for all grade levels. This threshold was different than already taught mathematics for third and sixth grade students. Kindergarteners being ‘Universal Novices’ could be the explanation for this disparity (Carey, 1985). For a kindergartener, everything in the academic world is new, even mathematics described as already taught, so it would make sense if the results matched new mathematics for other grade levels.

The already taught mathematics for third and sixth grade students showed a significant threshold, however the threshold went in a different direction. For this the lower spatial skill groups for both third and sixth grade student had less relation to mathematics, while the higher spatial group had a higher slope and relation to mathematics. My hypothesis for these results was that there would be no threshold, begging the question why the higher spatial skill group had a significantly higher slope. This could be due to some already taught mathematics questions requiring some spatial skills. While this knowledge may be largely procedural, difficult mathematics problems that are already taught could still require spatial skills.

Kindergarten

For kindergartners, when analyzing spatial thresholds for new mathematics, a significant threshold was found that matched Figure 5b (see Figure 10). This indicates that the lower spatial skill scoring students had a higher slope than the higher spatial skill scoring students. The Fisher's Z-score showed significant differences in strength of correlation between the two segments, with the lower segment having a significant correlation while the higher spatial skills group did not (see Table 4). These Fisher's Z-score results were further evidence for the threshold hypothesis. The already taught mathematics for kindergarteners had similar results that resulted in a significant threshold (see figure 11).

These results indicate that for both new and already taught mathematics students' spatial scores are strongly related up to a threshold. After spatial skill scores exceed that threshold, the relation between spatial skill scores and new and already taught mathematics significantly decreases in both slope and magnitude of correlation. This suggests that kindergarten students only need a certain level of spatial skills for both new and already taught mathematics, as additional spatial skills do not have as much of an impact on either new or already taught mathematics scores.

The results for new mathematics in kindergarteners suggest the threshold hypothesis applies to this area. Previous research confirms that students rely more heavily on spatial skill when performing new, unlearned, mathematics. However, already taught mathematics had a similar relation to spatial skills as did new mathematics. This result is contrary to previous research, and may be due to the age of the participants. Research has shown that younger students rely more on spatial skills to perform mathematics than do older students. Further, younger students have not had enough time to develop procedural knowledge of mathematics, and there is evidence that students with procedural knowledge of mathematics are less likely to rely on spatial skills when performing mathematics. In this interpretation, potentially all classroom mathematics is new to kindergarten students; therefore the distinction between new and already taught mathematics is negligible.

Third Grade

Third grade students also had significant spatial skill thresholds for both new and already taught mathematics (see Figures 12 and 13). The spatial threshold for new mathematics Davies test slope matched Figure 5b, suggesting that this paper's hypothesis is correct; however, the Fisher's Z-score did not provide additional evidence to bolster this study's hypothesis. The Davies test results indicated that the slope between mathematics and spatial scores significantly decreased for the higher scoring skills group (see Figure 12). The Fisher's Z-score outcome indicated that both high and low spatial skill scoring groups' correlations were significant, and that there was a significant difference between the correlations (see Table 4). Overall, third grade new mathematics did show a significant threshold, just not as strong of a threshold as appeared in kindergarten.

Already taught mathematics Davies test results matched Figure 5c (see Figure 13). This outcome means there was a significant change in slope, with the higher spatial skills scoring group having a higher slope than the lower scoring skills group. The Fisher's Z-score indicated that both high and low scoring skill groups' correlation is significant, and that there was no significant difference between the correlations (see Table 4).

New mathematics for third grade students showed that lower spatial skill scores and mathematics had a higher slope, and higher spatial skill scores had a significantly lower slope, matching Figure 5b (see Figure 12). This relation between new mathematics and spatial skill scores is due to the nature of spatial skills' use in new mathematics, while already taught mathematics relies more on procedural knowledge than spatial skills. These results indicate that the third grade students who are above the spatial threshold for new mathematics will not be as significantly impacted by spatial skills training as would the lower skills scoring group. The spatial skills threshold for third grade new mathematics could be due to students needing only a certain level of spatial skills to perform mathematics; so that for students above that threshold, the relation between mathematics performance and spatial skills decreases.

For already taught mathematics there is a lower slope between mathematics and spatial skill scores in the lower spatial skills group and a higher slope in the higher spatial skills group, matching Figure 5c (see Figure 13). In this study I hypothesized that already taught mathematics may not have a significant threshold, but these results state that there is a threshold going the other way. These results can be explained by some of the already taught problems still requiring spatial skills. While all of the problems in the already taught group should be easier for students because they can simply use procedural knowledge, some of the problems may be demanding enough that spatial skills are still required. Therefore, only students with a high level of spatial skills will be able to solve these mathematics problems. Since the majority of problems rely on procedural knowledge that the students already possess, some students in the lower spatial skill group will have a more consistent mathematics score with less of a relation with spatial skills as shown in this analysis.

Sixth Grade

Significant spatial thresholds were found for new and already taught mathematics with sixth grade students as well (see Figures 14 and 15). For new mathematics, Davies test results matched Figure 5b, suggesting that this paper's hypothesis is correct; however, the Fisher's Z-score did not provide evidence for this study's hypothesis. The Davies test indicated that the slope between mathematics and spatial skill scores significantly decreased for the higher scoring group. The Fisher's Z-score outcome showed that both high and low scoring groups' correlations were significant, and that there was a significant difference between the correlations. Overall, sixth grade new mathematics did show a significant threshold that was evident of this study's hypothesis, just not as strongly as that for the kindergarten age group, because for the kindergarten group, both the Davies test and the Fisher's Z-score were indicative of a threshold.

Already taught mathematics in sixth grade students' Davies test results matched Figure 5c (see Figure 15). This outcome means there was a significant change in slope, with the higher spatial skill scoring group having a higher slope than the lower scoring group. The Fisher's Z-score indicated that both the high and the low skill scoring groups' correlations of the two constructs were significant, and that there was no significant difference between the correlations (see Table 4).

New mathematics for sixth grade students showed that lower spatial skill scores and mathematics had a higher slope, and that the higher spatial skill score group had a significantly lower slope (see Figure 14). This relation between new mathematics and spatial skill scores is due to the nature of spatial skills' use in new mathematics, while already taught mathematics relies more on procedural knowledge than spatial skills. These results indicate that the sixth grade students who are above the spatial skill threshold for new mathematics will not be as significantly impacted by spatial skill training as the lower scoring group. The spatial skill threshold for new mathematics may be due to students needing a specific amount of spatial skills to be skilled at new mathematics, such that for students above that threshold the relation between spatial skill and mathematics decreases.

For already taught mathematics there is a lower slope between mathematics and spatial skill scores in the lower spatial skill group and a higher slope in the higher spatial skills group (see Figure 15). In this study I hypothesized that already taught mathematics may not have a significant threshold, but these results state there is a threshold going the other way. These results can be explained by some of the already taught problems still requiring spatial skills. While all of the problems in the already taught group should be easier for students because they can simply use procedural knowledge, some of the problems may be demanding enough such that spatial skills are still required. Therefore, only students with a high level of spatial skills will be able to solve these mathematics problems. Since the majority of problems rely on procedural knowledge that the students already possess, most students will have a more consistent mathematics score with less of a relation to spatial skills, as shown in this analysis. An alternative explanation for why new and already taught mathematics had different results for third and sixth grade students that different spatial skills are used when performing new, compared to already taught mathematics. A study by Mix et al (2016) performed an EFA and CFA found that already taught mathematics had a stronger relation to spatial visualization and new mathematics had a stronger relation to symbol dissemination. Thresholds for new and already taught mathematics may change if a threshold is looked for using specific spatial assessments.

Aim 4: Age Dependent Spatial Thresholds

The fourth research question asked if there were different thresholds for spatial skills depending on the students' ages. To probe for the existence of different age-based thresholds, composite mathematics and composite spatial scores were created for each grade level. This analysis was done because prior work suggests that the relation between spatial skills and mathematics changes depending on age (Holmes & Adams, 2006, Holmes et al., 2008, McKenzie et al., 2003, Rasmussen & Bisanz, 2005, Hecht et al., 2001, Noël et al., 2004 and Passolunghi et al., 2007). Previous research has indicated that older students use spatial skills less when performing mathematics. Based on this prior research I hypothesize that a significant spatial threshold will exist for all ages, but that the spatial threshold will change depending on grade level. Specifically, that older students will have a lower spatial threshold than the younger students.

However, my results did not fully match my hypothesis, as a significant spatial threshold was found only for third and sixth grade; in addition, sixth grade students had a higher spatial threshold than third grade students. When analyzing thresholds for kindergarten students, no significant threshold in terms of change in slope was found, and therefore no support for the Threshold Hypothesis for kindergarten students overall scores. The Davies test results matched Figure 5a; however, there was a significant threshold found in the Davies test that provided evidence for this study's hypothesis for both third and sixth grade students that matched Figure 5b (see Figure 17 and 18).

Third grade students' Davies test results demonstrated a spatial threshold that matched Figure 5b (see Figure 17). This indicates that students who are below the spatial skill threshold have higher correlations of mathematics and spatial skill than do students with higher spatial skills. The third grade students' Fisher's Z-score showed that both segments' correlation was significant, and that there was no significant change in correlation.

These same results, with a different threshold point, were found for sixth grade students (see Figure 18). These results indicate that third and sixth grade students only need a certain level of spatial skills to be skilled at mathematics, and any additional spatial skills would not be effective at improving students' mathematics scores. An alternative explanation for these results could be that students develop procedural knowledge, and rely less on spatial skills when solving mathematics problems.

For kindergarten students, a threshold was not found using the Davies test (see Figure 16). Kindergarteners' analysis matched Figure 5a, which states no significant change in slope was found between the two groups. However, the Fisher's Z-score for overall kindergarteners showed that both high and low spatial skill groups have a significant correlation, and that there was a significant difference in correlation with the lower spatial skill group displaying a weaker correlation (see Table 4).

These results provided mixed evidence for the age dependent spatial skill threshold hypothesis in this paper. The lack of a significant threshold found for kindergarten students goes against the hypothesis that stated spatial skills will become less related to mathematics once kindergarteners reach a certain level of spatial skill. This research indicates that the strength of correlation and slope of mathematics and spatial scores is maintained for kindergarteners. Previous research found a stronger relationship between spatial and mathematics skills for younger students. This could explain the lack of significant change in slope, because kindergarteners rely heavily on spatial skills to perform all mathematics problems. Therefore, kindergarten students need spatial skills to solve all mathematics problems, even ones they have encountered before. Such a reliance on spatial skills may also be due to kindergarteners not having already internalized the procedural knowledge required to solve mathematics problems, making all problems reliant on spatial skills. The lack of internalizing procedural knowledge may be because they have not had the time to learn the procedural knowledge of kindergarten level mathematics, or that students that age lack the cognitive abilities to learn procedural mathematics knowledge. Thus, older students would draw from a larger procedural knowledge bank of approaches to math problems than kindergarten students would. Another explanation could be that students who are performing new mathematics rely more on spatial skills than students performing already taught mathematics; all mathematics in an educational setting is new to kindergarteners and will require spatial skills to perform.

The threshold hypothesis was found to be applied to spatial and mathematics skills in both third and sixth grade students. Third grade students had a lower threshold than sixth grade students. This difference in threshold between third and sixth grade students means that sixth grade students needed to reach a higher level of spatial skills than third grade students to be in the ‘higher’ spatial group. These results could be due to the fact that sixth grade students were assessed in areas such as geometry, graphs, and algebra which use more spatial skills than the multiplication, division, and missing terms problems that third grade students were assessed on.

Conclusions

This study revealed spatial thresholds as predicted by the Threshold Hypothesis. Thus, there is reason to conclude that superior spatial skills are related to superior mathematics performance, but only up to a certain level of proficiency. Furthermore, these spatial skills thresholds varied, depending on the student's age and on the type of mathematics problem. Sixth grade students had a higher skills spatial threshold than third grade students, with kindergarteners having no significant threshold. My initial hypothesis was that younger students would have a higher spatial threshold; however my results did not support this. These results may be due to sixth grade students being assessed on more complex mathematical concepts such as geometry and knowledge of graphs which would require a higher level of spatial skills, while younger students were assessed on simpler mathematics.

Third and sixth grade students showed a significant threshold that supported this study's hypothesis, while kindergarten students had a strongly correlated spatial and mathematics score, but no significant threshold. This indicates that training for all kindergarten students in spatial skills will transfer to mathematics, while spatial training for third and sixth grade students would be more effective for students below the spatial threshold. The lack of a significant threshold in kindergarteners relative to the other grades may be due to kindergartener's heavy reliance on spatial skills to perform mathematics or to their lack of procedural knowledge.

When comparing new and already taught mathematics in Aim 3, a stark contrast was shown between kindergarteners and the other grade levels. For both new and already taught mathematics in kindergarteners, a threshold was found that suggested students only needed a certain level of spatial skill and any more would not have as big of an impact on spatial scores. A potential explanation for kindergarteners' unique results is that kindergarteners are 'Universal

Novices', which means that everything they learn in school is new to them, even mathematics problems that were labeled already taught. Therefore, even already taught mathematics for kindergarteners would have results that matched new mathematics for all grade levels.

The main question that drove this study was, why do spatial training experiments that are meant to transfer to mathematics have mixed results? This study may provide an explanation as to why prior research has shown mixed results with respect to training spatial skills and its transfer to mathematics. Because of the strong relation between spatial skills and mathematics, researchers have hypothesized that training in spatial skills could improve mathematics scores. These studies do not take into consideration that students may only need a certain level of spatial skills to perform mathematics, which could skew results. Students who are above the spatial threshold in these training studies would not show as significant an impact on mathematics scores as students who are below the spatial skill threshold. I hypothesize that if researchers trained students with spatial scores who were below the spatial threshold, significant increases in mathematics scores would be found.

In this paper I have explained the existence of a spatial skill threshold for mathematics in two possible ways for most circumstances. One possibility is that students learn procedural knowledge after which they no longer rely on spatial skills as much. The higher spatial skill group students may be better able to learn and deploy procedural knowledge for mathematics, such that they need not revert to spatial skills to conceptualize and resolve problems. The other explanation for a spatial threshold is that students only need a certain amount of spatial skill to perform mathematics, and that once students exceed that threshold further training is no longer beneficial. This interpretation of the results harkens back to the computer analogy made earlier in this paper. This analogy stated that spatial skills are similar to hardware, while mathematics knowledge is similar to software. If a computer's hardware can already run certain software, then upgrading the hardware will not improve the performance of the software.

In terms of a threshold for spatial skills in mathematics, regardless of explanation, this study's findings indicate that when students do not have the required level of spatial skills, improvement in spatial skills will improve mathematics skills as well. However, when spatial skills are above that threshold, any further improvement will not have as significant an impact on mathematics scores. The two possible explanations for this study's results — that students develop procedural knowledge, and that students only need a certain level of spatial skills to perform mathematics — are not mutually exclusive. In fact, I suggest that both these explanations obtained in most of the results that were found to support this study's hypotheses. Nonetheless, the notion of a spatial skills threshold holds greater explanatory power.

However learned procedural knowledge cannot explain the thresholds found for new mathematics, in Aim3, because a teacher has not had the opportunity to teach students this mathematical procedural knowledge. Therefore, procedural knowledge is not a valid explanation in the new mathematics analysis; leaving the explanation that a student only needs a certain level of spatial skills to be skilled in mathematics. As stated before, I do believe that both procedural knowledge and students only needing a certain level of spatial skills are probable causes for this study's results, however this additional evidence for students needing a certain level of spatial skills is important to note when discussing the effect of spatial training on mathematics scores.

Bidirectionality of Spatial and Mathematics Skills

Due to the strong relation between spatial and mathematics skills, one could suggest spatial and mathematics problems use the same skills to solve problems; however, the current data and prior research has not shown support for this idea. While the correlation between spatial and mathematics is clearly bidirectional, no researchers have shown that training students in mathematics will improve the students' spatial skills. In addition, there are several studies that show spatial skills predict future mathematics performance, but no research has been published showing that mathematics predicts spatial scores (Alloway & Passolunghi, 2011; Gathercole & Pickering, 2000; Johnson, 1998; Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki, 2003; Lachance & Mazzocco, 2006; Markey, 2010; Mazzocco & Myers, 2003; Mix et. al., 2016). This is not to say that training mathematics would not improve spatial performance, or that mathematics cannot predict spatial performance; simply that researchers have looked at this relationship going one way. Researchers' choice of analyzing spatial and mathematics skills relation going one way reflects our understanding of the spatial and mathematics relation; indicating that spatial skills are a more general skill used to perform mathematics, while mathematics is a more specific skill that will not transfer as well to spatial assessments.

Because the present study was motivated by interest in whether spatial training can improve mathematics performance, I also focused on thresholds in spatial skill. However, I also explored whether there were significant thresholds in the other direction of causality. That is, if one wanted to improve spatial skill through mathematics training, does the relation between spatial skill and mathematics change at a certain level of mathematics proficiency. The details of these analyses are beyond the scope of this dissertation, but it is interesting to note that no significant mathematics threshold were found when analyzing mathematics as a whole and based on grade level, suggesting that this relation is not bidirectional when it comes to the Threshold Hypothesis.

Implications

This study aimed to explore the threshold hypothesis as a potential causal mechanism by which spatial skill is related to mathematics ability in order to provide evidence to explain why spatial training does not always result in subsequent improvement in mathematics scores. This study has provided a possible explanation for why spatial training studies in the past have not transferred to mathematics. This paper provides the groundwork for creating targeted spatial training, so training can focus on students who have not hit the spatial threshold needed to be skilled at mathematics. By adapting spatial training to target students below the spatial threshold, the spatial training will be more likely to transfer to mathematics skills. The results of this study may alternatively suggest that if a student does not have a deficit in spatial skills, their low mathematics score could be a result of a different problem. Thus, educators should not seek a one-size-fits-all solution when it comes to deficits in mathematics or spatial ability; rather, they should recognize the individual differences among their students and target instruction to address these distinctions.

In addition, this could apply to areas other than mathematics. Researchers have indicated that spatial skill deficits may provide a barrier for entry into STEM (Science, Technology, Engineering, and Mathematics) degrees (Uttal & Cohen, 2012). It is likely that the relationship of spatial skill with science, technology, and engineering is similar to the relationship of spatial skill with mathematics, so there could be a training effect by which spatial training would improve science, technology, and engineering scores. The threshold hypothesis in terms of a spatial threshold applied to science, technology, and engineering could help students overcome the barriers to entering the STEM fields. Other researchers have also suggested that spatial skills are needed for creativity; a similar threshold may exist for spatial skills and creativity. Creativity is used in many fields, and if spatial skills are the engine that drives creativity, it would be important to teach students spatial skills to achieve even higher levels of performance.

Limitations

A limitation is that in some analyses, the threshold for spatial skills to perform mathematics is far from the mean. This could be due to too many spatial questions being too easy or too difficult for specific populations in each analysis. This may have occurred because there has been less testing and standardization than in other areas that have tested for threshold, such as IQ. A floor or ceiling for the data could negatively impact the statistical analysis due to an imbalance in the number of participants in the low and high spatial skill group comparisons.

Another limitation addresses the lack of significance for Aim 2 that looked at specific spatial assessment's thresholds when it came to mathematics. This analysis could have been flawed due to the individual assessment not being robust enough for the analysis, due to too few problems in each assessment. Having too few problems may have caused a significant change if students missed even a single problem, as some assessments had as few as four problems. A student missing a single problem could cause a 25% change in their scores. In addition, too few questions could prevent the construct from being fully tapped, because students would more easily hit floor or ceiling.

Additionally I analyzed for a spatial threshold instead of other general skills such as working memory. It is possible that working memory or other general skills could also has a threshold when it comes to mathematics, however spatial skills have been shown to have a stronger relation with mathematics (Reuhkla, 2001). In the future it would be beneficial to perform a segmented regression analysis on mathematics and other general skills, to determine if a Threshold exists with mathematics and other skills; however, this data set was limited to spatial and mathematics assessments.

APPENDIX

Table 1:

Mathematics and Spatial Assessments Based on Grade Level

Kindergarten Mathematical Assessments	**3rd Grade Mathematics Assessments**	**6th Grade Mathematics Assessments**
Place Value	Place Value	Place Value
Word Problems	Word Problems	Missing Term/Algebra
Computational Thinking	Computational Thinking	Word Problems
Missing Term/Algebra	Missing Term/Algebra	Computational Thinking
Number Line Estimation	Number Line Estimation	Multi-Step Calculation
Fractions	Fractions	Notational Spacing
Proportional Reasoning	Multiplication	Number Line Estimation
Calculation	Calculation	Fractions
		Chart, Tables, and Graphs
		Geometry
		Proportional Reasoning

Kindergarten Spatial Assessments	**3rd Grade Spatial Assessments**	**6th Grade Spatial Assessments**
Mental Rotation	Mental Rotation	Mental Rotation
Visual Spatial Working Memory	Visual Spatial Working Memory	Visual Spatial Working Memory
Test of Visual Motor Integration	Test of Visual Motor Integration	Test of Visual Motor Integration
Block Design	Block Design	Block Design
Map Reading	Map Reading	Map Reading
Perspective Taking	Perspective Taking	Perspective Taking

Table 2:

Correlations between Mathematics and Spatial Scores

	Correlation Low	Correlation High	Fisher's Z-score	P-value
Whole	.415*	.564**	-1.09	.138
Block Design	.56**	.44**	2.07	.039
Perspective Taking	-1.0	.34**	-8.3	0
Mental Rotation	.274**	.399**	-2.41	.016
New Kindergarten	.789**	.319	2.07	.019
Already Taught Kindergarten	.602**	-.404	4.56	0
New Third Grade	.728**	.799**	-1.91	.028
Already Taught Third Grade	.844**	.845**	-.04	.484
New Sixth Grade	.655**	.795**	-3.32	.0005
Already Taught Sixth Grade	.667**	.610**	1.02	.154
Kindergarten	.318**	.525**	-2.54	.0055
Third Grade	.534**	.557**	.86	.195
Sixth Grade	.408**	.315**	1.22	.111

** p < .01, * p < .05

Table 3:

Means and Standard Deviations of Low and High Spatial Groups

	Spatial Mean	Spatial Standard Deviation	Mathematics Mean	Mathematics Standard Deviation
Whole: Low Spatial Group	-1.482	.160	-1.176	.728
Whole: High Spatial Group	.026	.715	.032	.677
Block Design: Low Spatial Group	-1.479	.318	-.564	.685
Block Design: High Spatial Group	.202	.885	.078	.709
Perspective Taking: Low Spatial Group	-2.736	.174	-.29	.644
Perspective Taking: High Spatial Group	.011	.875	.002	.737
Mental Rotation: Low Spatial Group	-.44	.652	-.142	.691
Mental Rotation: High Spatial Group	1.372	.551	.448	.692

Table 3 (cont'd)

	Spatial Mean	Spatial Standard Deviation	Mathematics Mean	Mathematics Standard Deviation
New Kindergarten: Low Spatial Group	-.034	.322	-.045	.687
New Kindergarten: High Spatial Group	1.578	.286	2.184	.389
Already Taught Kindergarten: Low Spatial Group	-.039	.417	-.029	.726
Already Taught Kindergarten: High Spatial Group	1.147	.160	.762	1.020
New Third Grade: Low Spatial Group	-.526	.262	-.673	.538
New Third Grade: High Spatial Group	.319	.332	.411	.475
Already Taught Third Grade: Low Spatial Group	-.630	.414	.850	.513
Already Taught Third Grade: High Spatial Group	.247	.247	.334	.484

Table 3 (cont'd)

	Spatial Mean	Spatial Standard Deviation	Mathematics Mean	Mathematics Standard Deviation
New Sixth Grade: Low Spatial Group	-.427	.201	-.642	.547
New Sixth Grade: High Spatial Group	.253	.319	.382	.530
Already Taught Sixth Grade: Low Spatial Group	-.173	.420	-.295	.620
Already Taught Sixth Grade: High Spatial Group	.383	.065	.644	.502
Kindergarten: Low Spatial Group	-.303	.399	-.188	.639
Kindergarten: High Spatial Group	.839	.482	.523	.799
Third Grade: Low Spatial Group	-1.473	.183	-1.143	.794
Third Grade: High Spatial Group	.065	.650	.052	.679
Sixth Grade: Low Spatial Group	-.566	.376	-.358	.652
Sixth Grade: High Spatial Group	.671	.466	.427	.573

Table 4:

Threshold Hypothesis Results

	Davies Test Outcome Matches Figure	Additional Evidence from Fisher's z-score	Evidence of a Threshold as Described by the Hypothesis
Whole	5b	No	Yes
Block Design	5c	No	No
Perspective Taking	5c	No	No
Mental Rotation	5c	No	No
New Kindergarten	5b	Yes	Yes
Already Taught Kindergarten	5b	Yes	Yes
New Third Grade	5b	No	Yes
Already Taught Third Grade	5c	No	No
New Sixth Grade	5b	No	Yes
Already Taught Sixth Grade	5c	No	No
Kindergarten	5a	No	No
Third Grade	5b	No	Yes
Sixth Grade	5b	No	Yes

Table 5:

Mathematics Problems Split into New and Already Taught

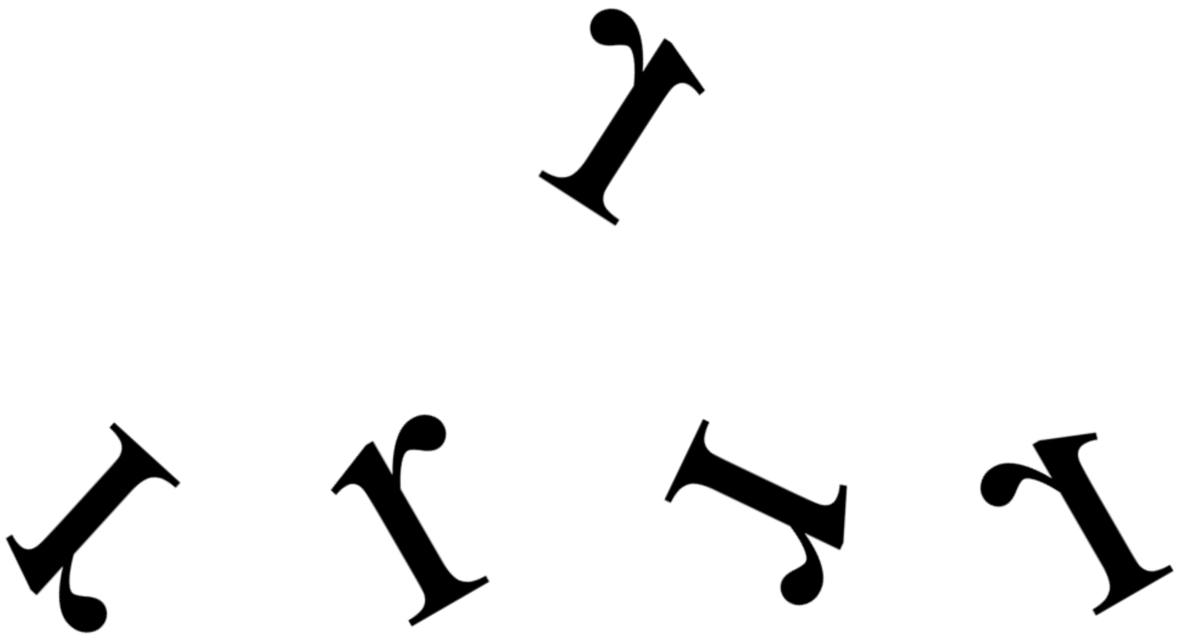
	Kindergarten	3rd	6th
Already Learnt	**Place Value:** 13 **Number Line:** 4 **Calculation:** Single digit addition and subtraction W**ord Problem:** TEMA 16 and 26	**Place value:** all **Number Line 0-1000:** all **Calculation**: Addition and subtraction **Missing term**: Addition and subtraction **Word problem**: TEMA 16, 17, 26, 32, 34, 52, 65, 67, 72	**Number line:** 0-100k, 0-1 **CMAT Problem solving:** 1-17 **CMAT Geometry**: 1-7 **Calculation** **Fraction** **CMAT Charts, tables, and graphs** **Rational number:** 4,7,11
New	**Place Value:** The rest of PV problems **Calculation**: Multidigit calculation **Word Problem:** TEMA 17, 25, 32, 34, 52, 57, 64, 65, 67, and 72. **Number Line:** except NL4 **Missing term**	**Fraction** **Multiplication** **Division** **Word Problem**: TEMA 25, 57, 64	**CMAT Algebra** **Problem solving**: 18-24 **CMAT Geometry:** 8-24 **CMAT other Rational Number items**

Table 6:

Comparison of Segmented Regression and Spline Analysis Fit

	Segmented	**Spline**	**Steiger's Z-score**	**P-value**
Whole	.3535	.3514	.068	.9459
Block Design	.2394	.2390	.011	.9914
Perspective Taking	.1631	.1612	.06	.9520
Mental Rotation	.1964	.1971	-0.021	.9835
New Kindergarten	.6876	.6929	-0.199	.8420
Already Taught Kindergarten	.3754	.3645	.272	.7857
New Third Grade	.8062	.8084	-0.151	.8801
Already Taught Third Grade	.8683	.8706	-0.198	.8430
New Sixth Grade	.7591	.7596	-0.029	.9766
Already Taught Sixth Grade	.6327	.6120	.954	.3401
Kindergarten	.3097	.3061	.095	.9244
Third Grade	.3856	.3852	.009	.9924
Sixth Grade	.3858	.3850	.022	.9825

Figure 1: Mental rotation task examples



1.a

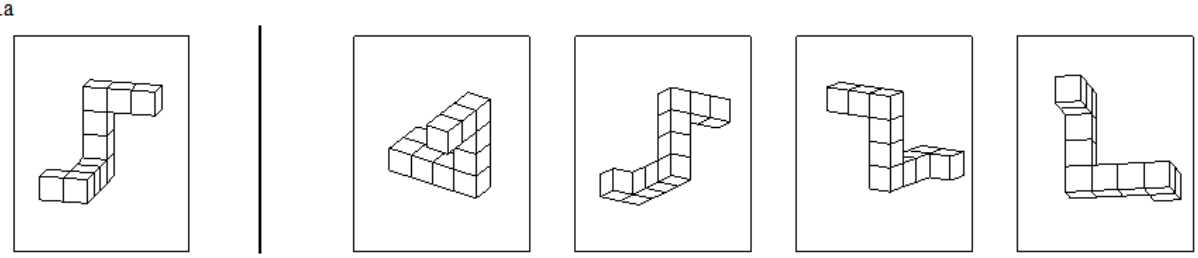


Figure 2: Map reading task

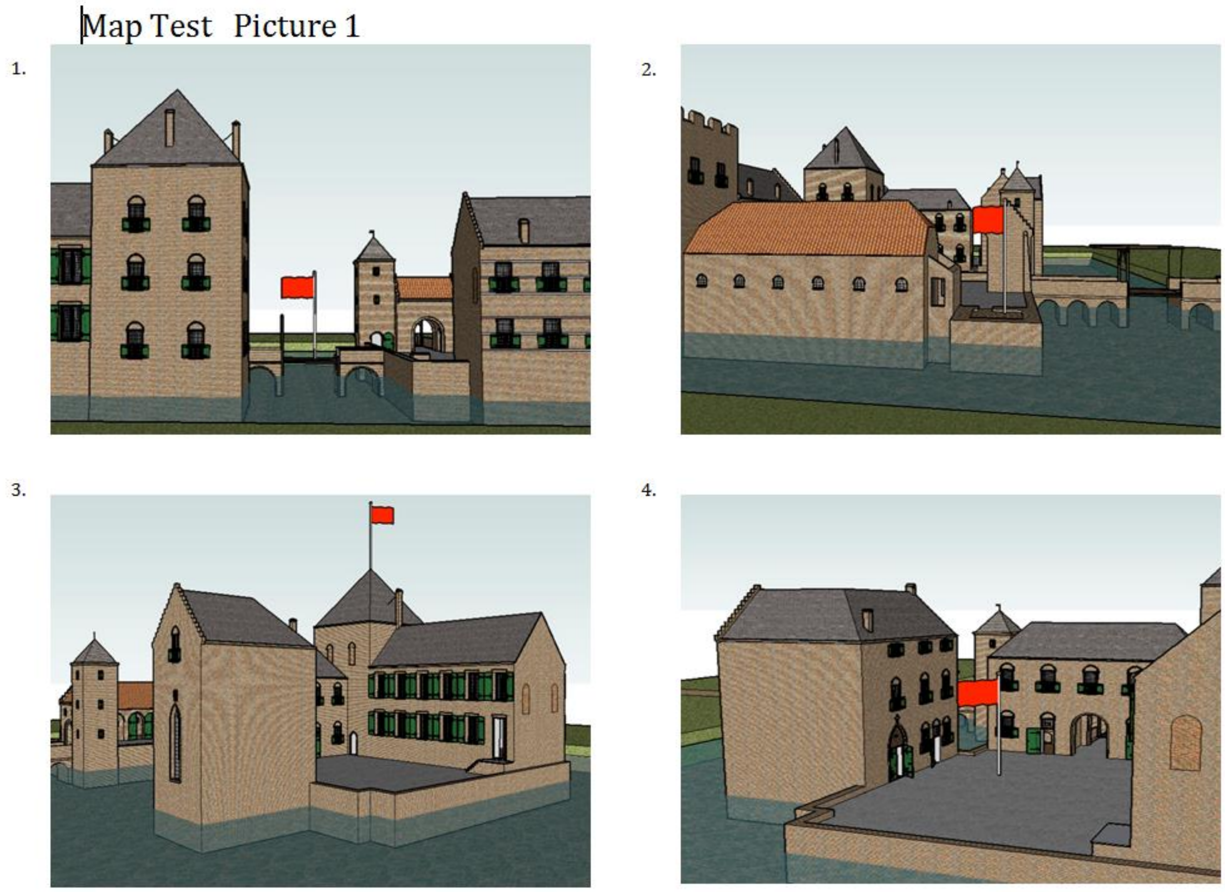


Figure 3: Kindergarten perspective taking task example

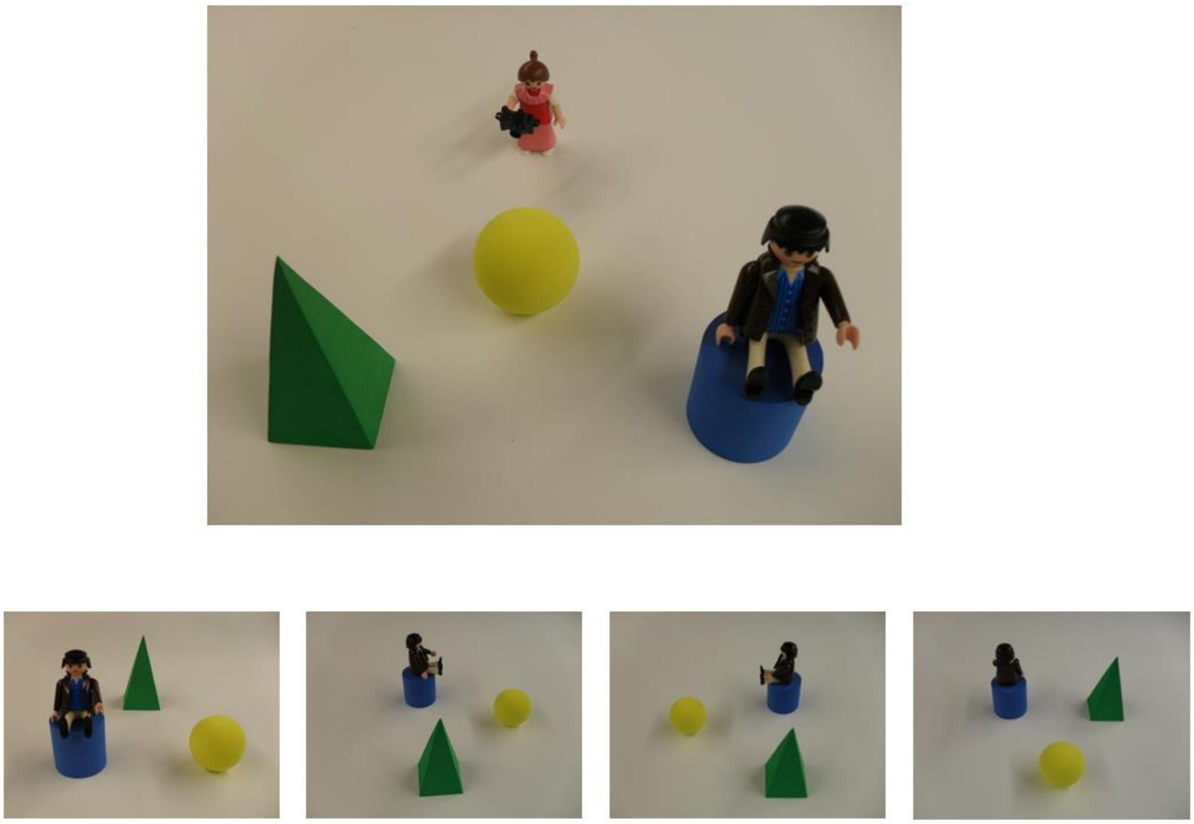


Figure 4: Third and sixth grade perspective taking task example

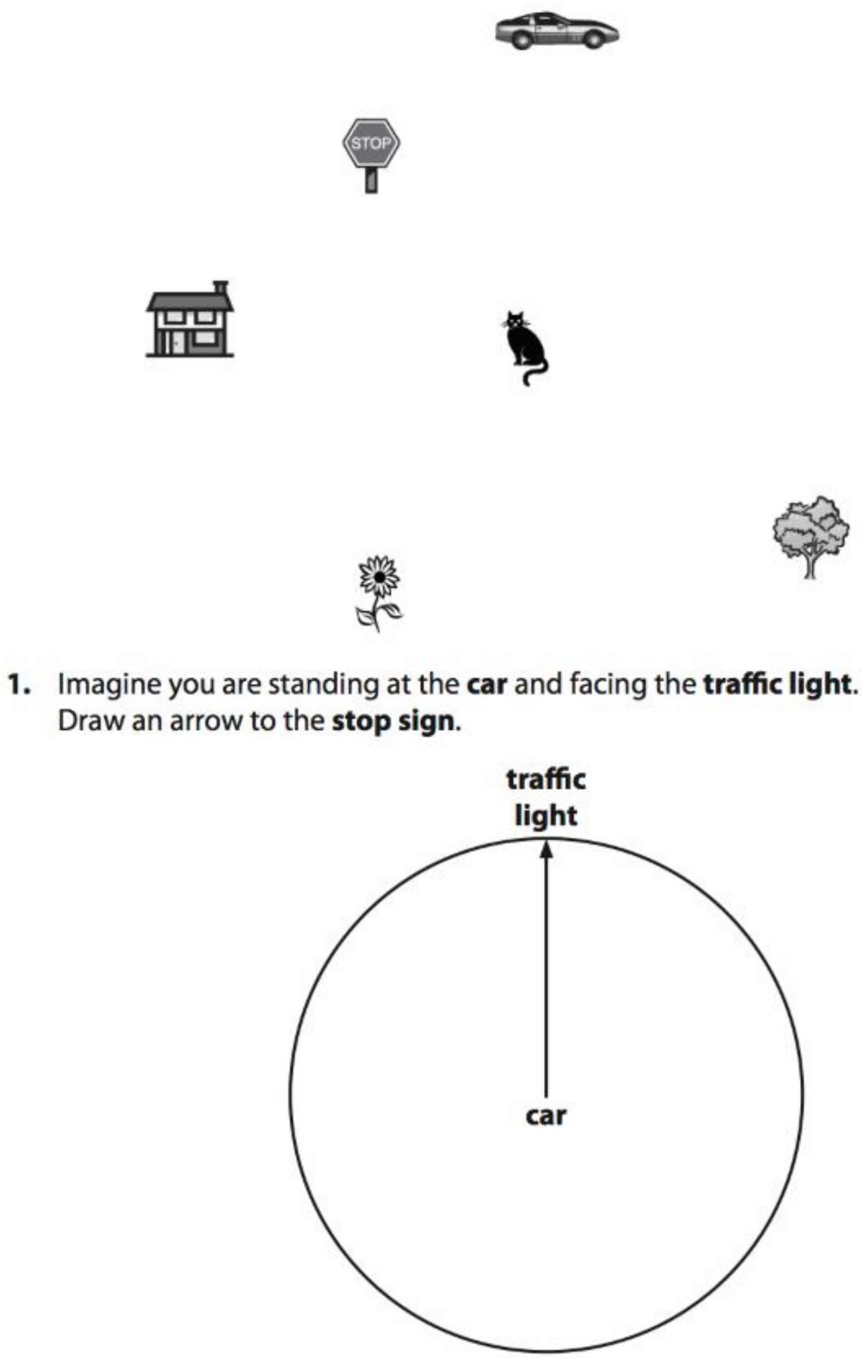


Figure 5: Overall Spatial Threshold for Overall Mathematics Model Graphs

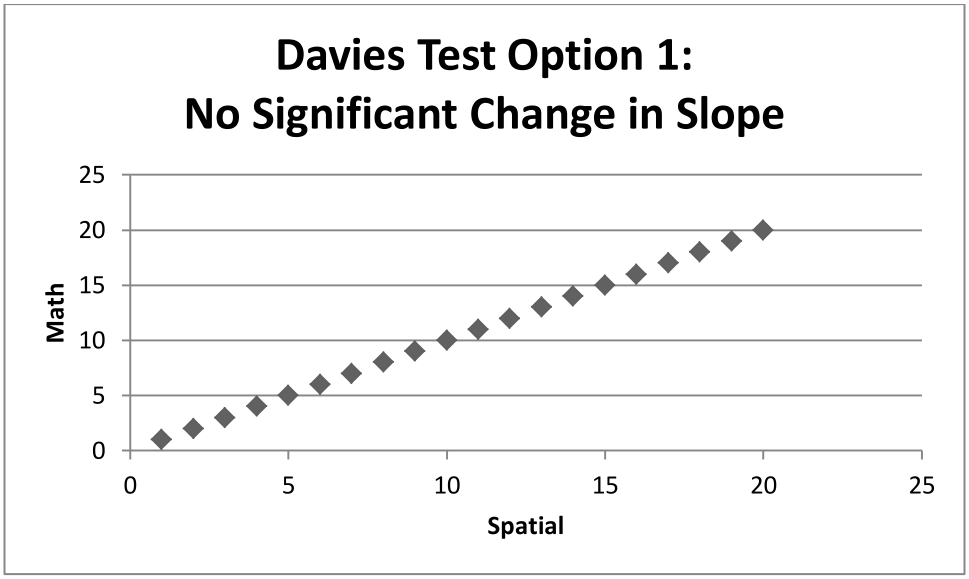


Figure 5b:

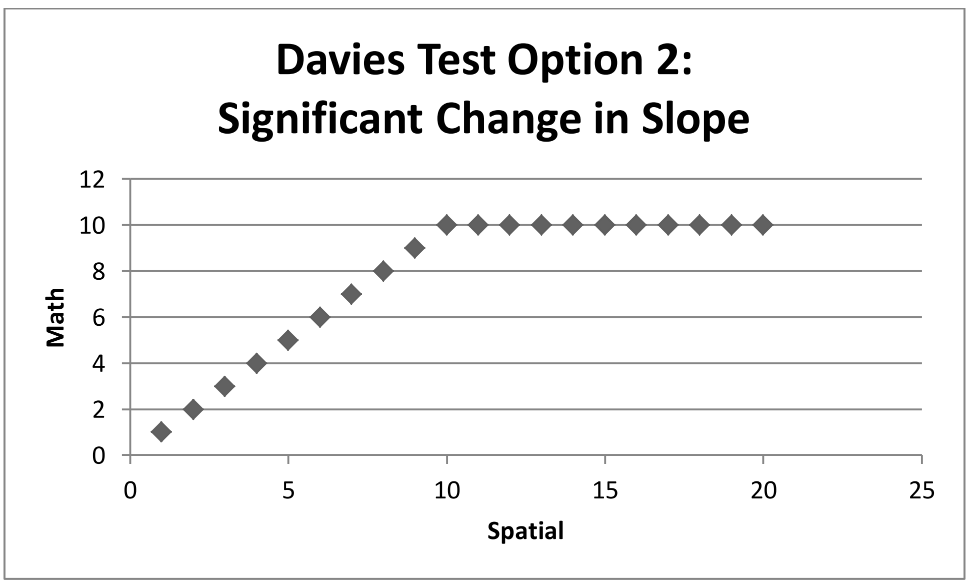


Figure 5 (cont'd)

Figure 5c:

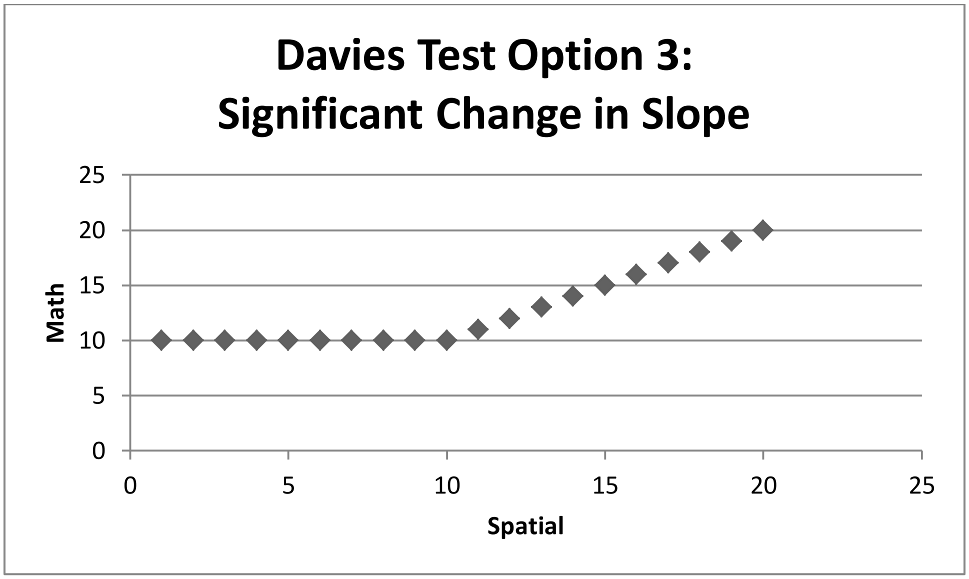


Figure 6: Overall Spatial Threshold for Overall Mathematics

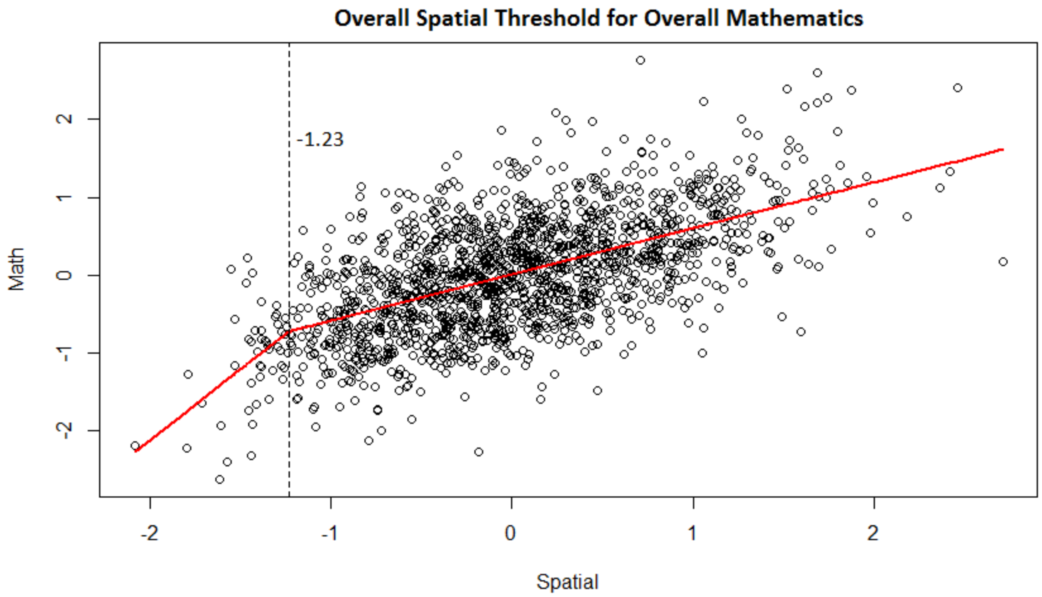


Figure 7: Block Design Threshold for Overall Mathematics

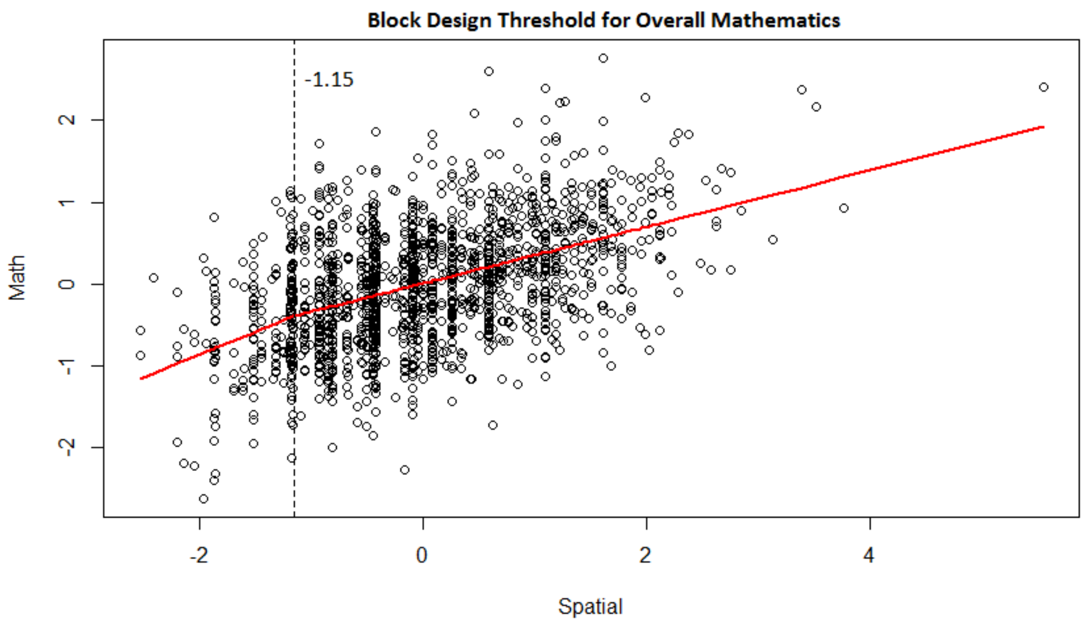


Figure 8: Perspective Taking Threshold for Overall Mathematics

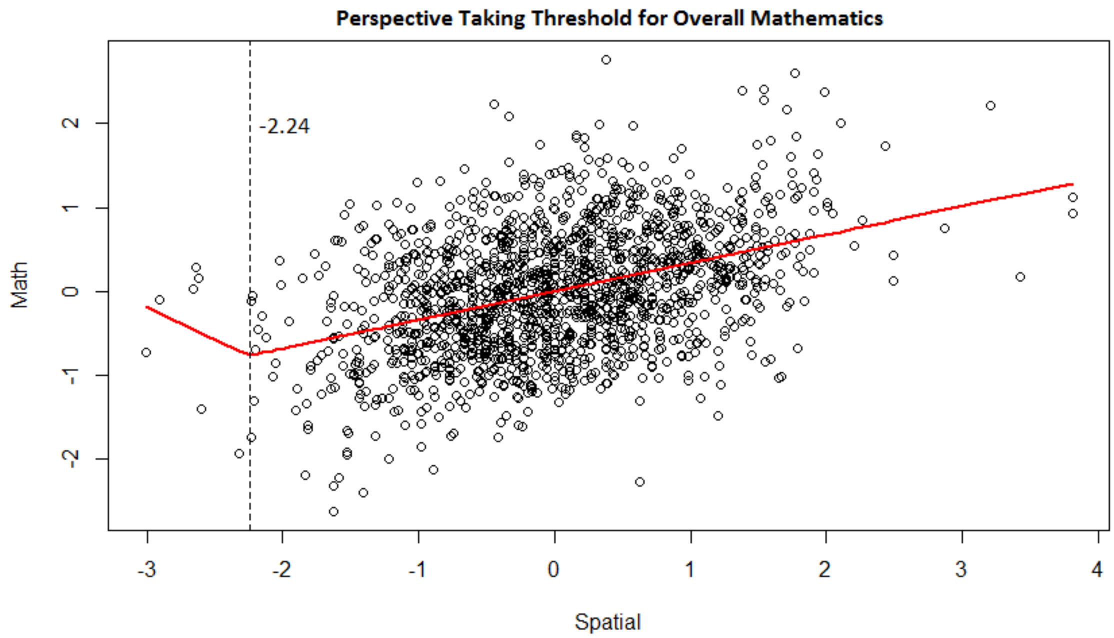


Figure 9: Mental Rotation Threshold for Overall Mathematics

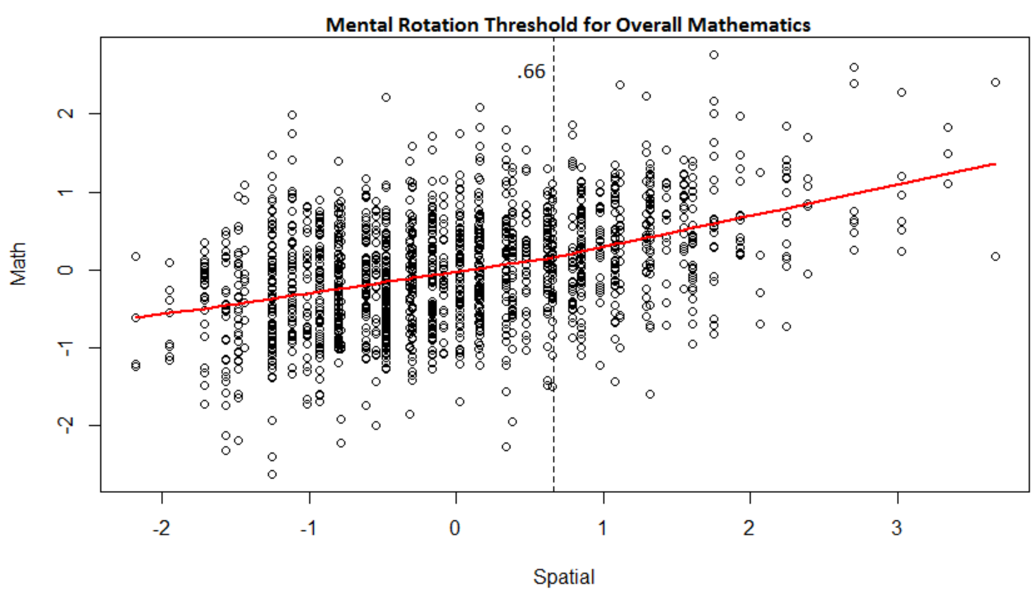


Figure 10: Spatial Threshold for New Mathematics in Kindergarteners

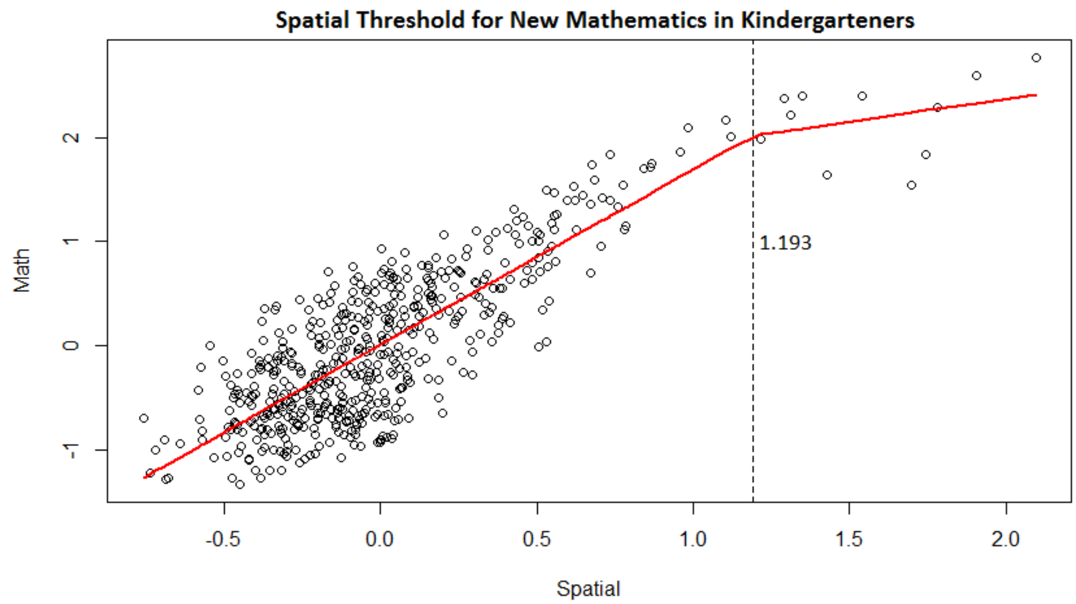


Figure 11: Spatial Threshold for Old Mathematics in Kindergarteners

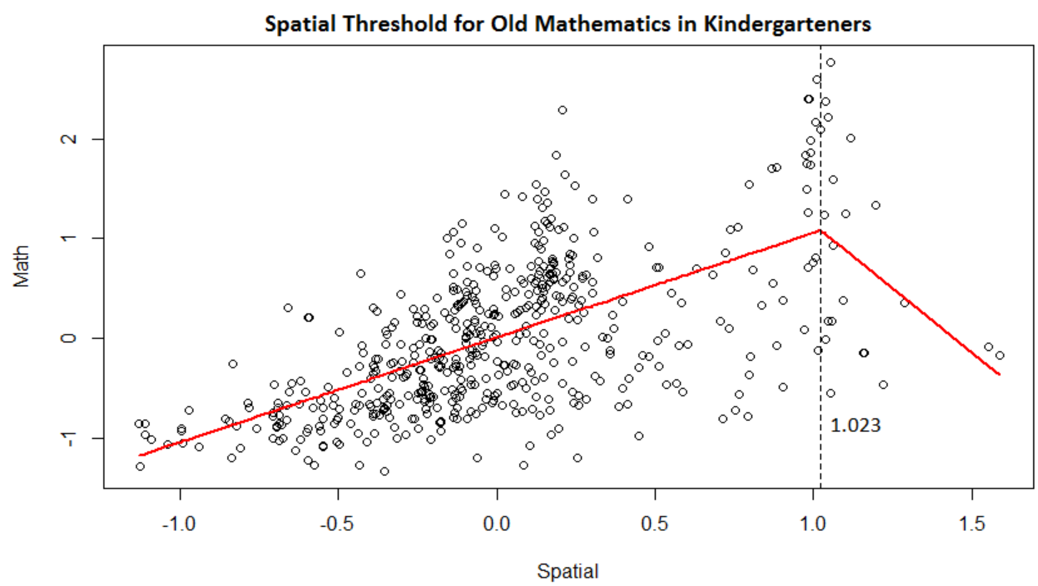


Figure 12: Spatial Threshold for New Mathematics in Third Graders

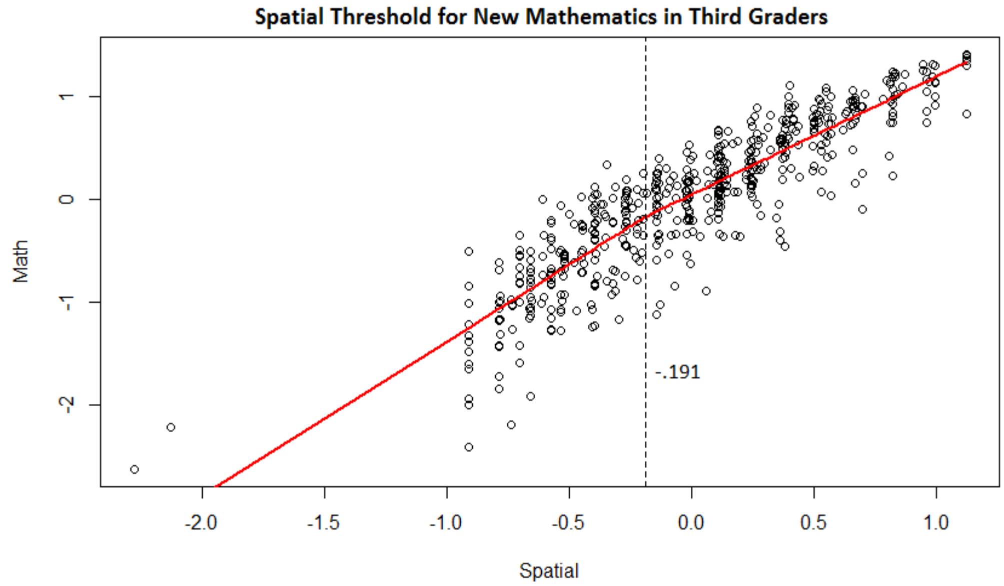


Figure 13: Spatial Threshold for Old Mathematics in Third Graders

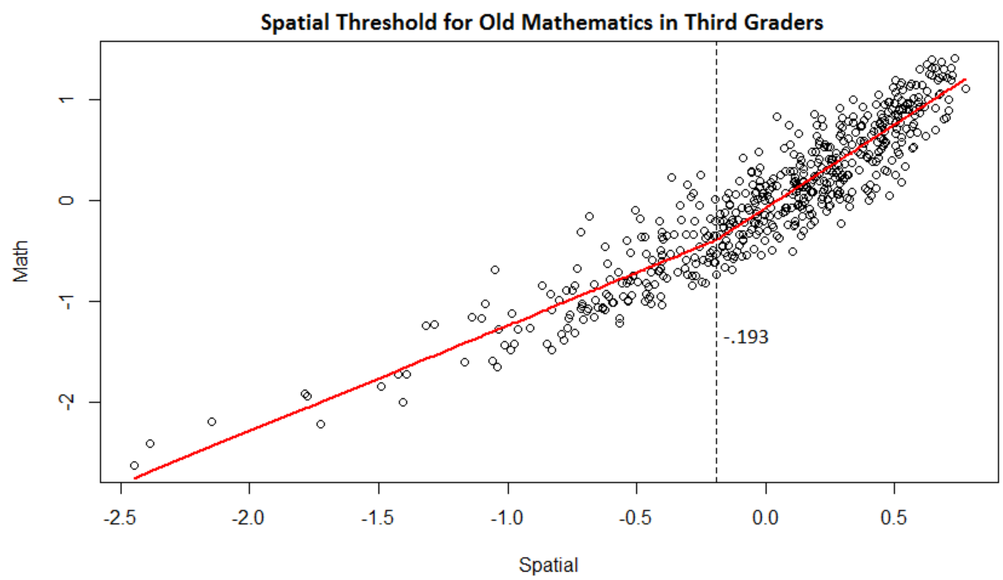


Figure 14: Spatial Threshold for New Mathematics in Sixth Graders

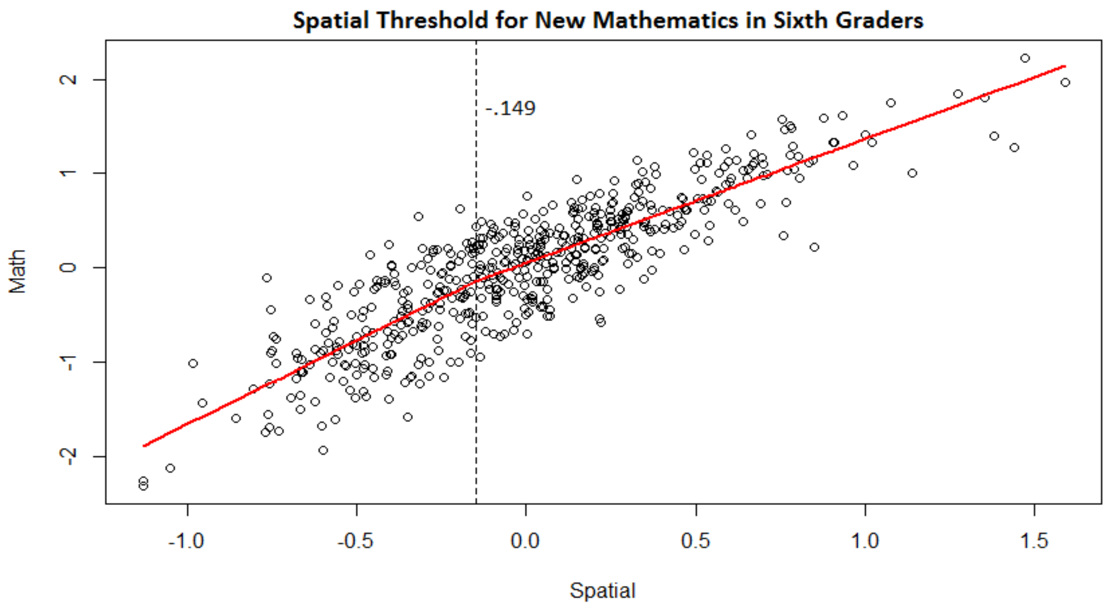


Figure 15: Spatial Threshold for Old Mathematics in Sixth Graders

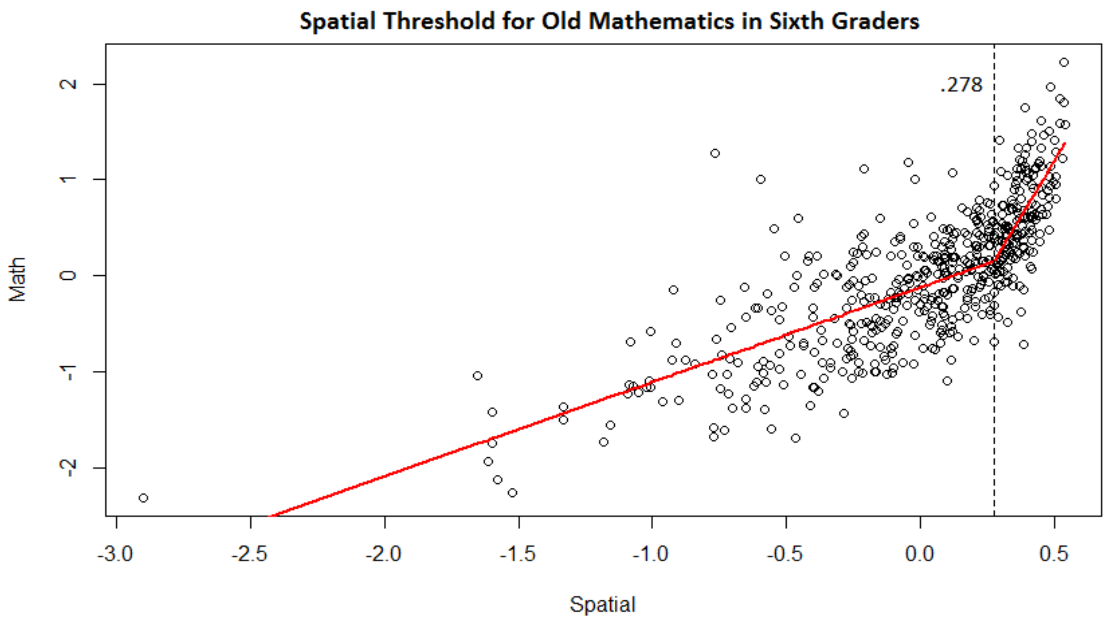


Figure 16: Spatial Threshold for Mathematics in Kindergarteners

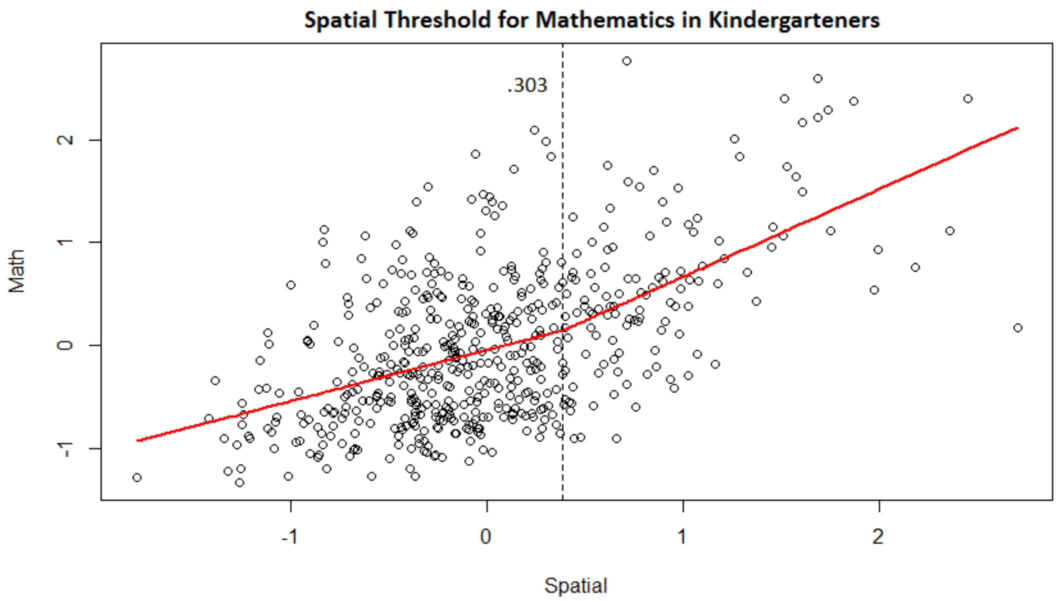


Figure 17: Spatial Threshold for Mathematics in Third Graders

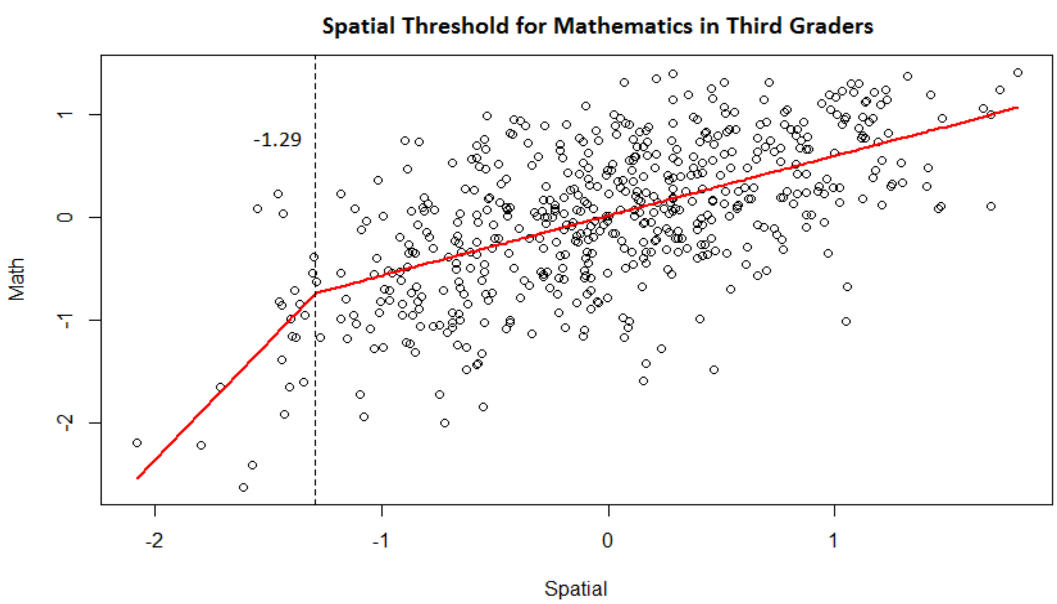
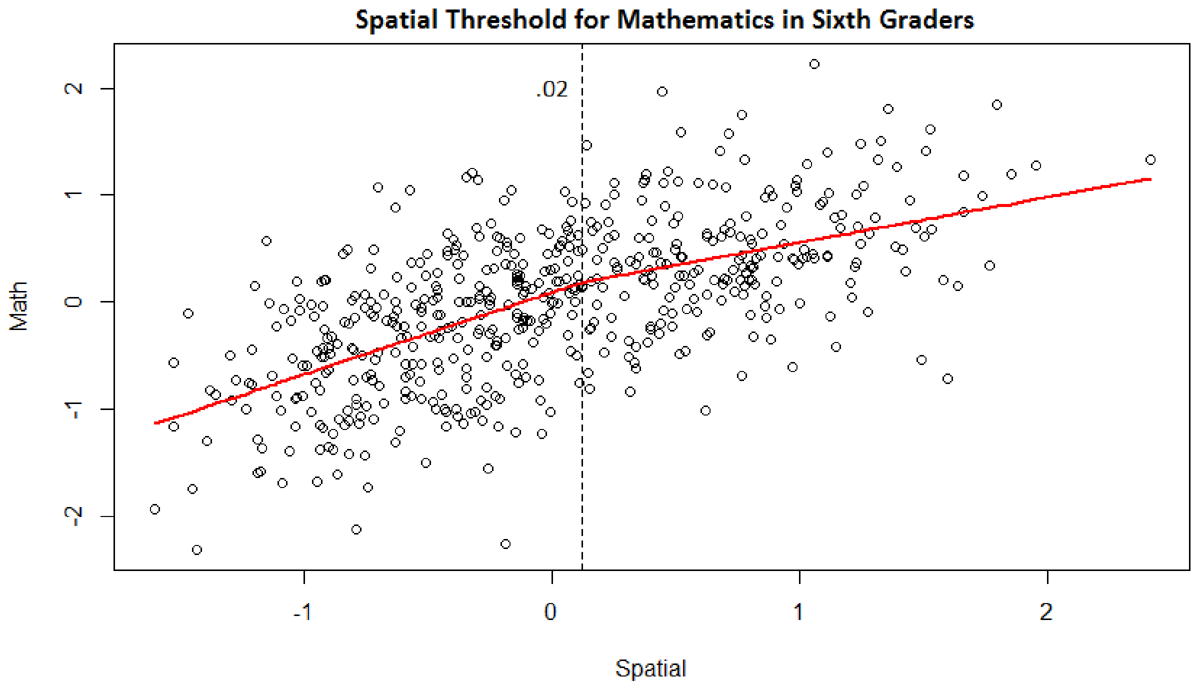


Figure 18: Spatial Threshold for Mathematics in Sixth Graders



REFERENCES

REFERENCES

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, 21, 133 – 137.

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex roles*, *20*(5-6), 327-344.

Beery, K. E. (2010). Administration, scoring and teaching manual for the Beery-VMI.

Bishop, A. J. (1980). Spatial abilities and mathematics education—A review. *Educational Studies in Mathematics*, *11*(3), 257-269.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental psychology*, *42*(1), 189.

Booth, R. D., & Thomas, M. O. (1999). Visualization in mathematics learning: Arithmetic problem-solving and student difficulties. *The Journal of Mathematical Behavior*, *18*(2), 169-190.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental neuropsychology*, *33*(3), 205-228.

Carey, S. (1985). Are children fundamentally different kinds of thinkers and learners than adults. *Thinking and learning skills*, *2*, 485-517.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, *15*(1), 2-11.

Clements, K. (1981). Visual imagery and school mathematics. *For the Learning of Mathematics*, *2*(2), 2-9.

De Lisi, R., & Cammarano, D. M. (1996). Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior*, *12*(3), 351-361.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of experimental child psychology*, *103*(2), 186-201.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex*, *22*(5), 1078-1085.

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental psychology*, *42*(6), 1259.

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of experimental child psychology*, *123*, 53-72.

Ferrini-Mundy, J. (1987). Spatial training for calculus students: Sex differences in achievement and in visualization ability. *Journal for Research in Mathematics Education,* 126-140.

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends in cognitive sciences*, *18*(10), 536-542.

Fuchs-Beauchamp, K. D., Karnes, M. B., & Johnson, L. J. (1993). Creativity and intelligence in preschoolers. *Gifted Child Quarterly*, *37*(3), 113-117.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology*, *70*(2), 177-194.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Developmental psychology*, *47*(6), 1539.

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.

Ginsburg, H., & Baroody, A. (2003). TEMA-3 Examiners manual.

Guilford, J. P. (1967). The nature of human intelligence.

Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children’s spatial and mathematics performance: A randomized controlled study. *Trends in Neuroscience and Education*, *4*(3), 60-68.

Heathcote, D. (1994). The role of visuo-spatial working memory in the mental addition of multi-digit addends. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*.

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology*, *79*(2), 192-227.

Hegarty, M. (2010). Psychology of Learning and Motivation.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of educational psychology*, *91*(4), 684.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175-191.

Heil, M., Rösler, F., Link, M., & Bajric, J. (1998). What is improved if a mental rotation task is repeated–the efficiency of memory access, or the speed of a transformation routine?. *Psychological Research*, *61*(2), 99-106.

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*(3), 339-366.

Holmes, J., Adams, J. W., & Hamilton, C. J. (2008). The relationship between visuospatial sketchpad capacity and children's mathematical skills. *European Journal of Cognitive Psychology*, *20*(2), 272-289.

Hresko, W., Schlieve, P., Herron, S., Swain, C., & Sherbenou, R. (2003). Comprehensive Mathematical Abilities Test (CMAT). *Austin: PRO-ED*.

Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education*, *86*(2), 151-158.

Hubber, P. J., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *The Quarterly Journal of Experimental Psychology*, *67*(5), 936-954.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. C. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence*, *41*(4), 212-221.

Johnson, E. S. (1998). An exploration of the relation between mathematics achievement and factors of intelligence. (Order No. AAM9828320, *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 724. Kail, R. (1986). The impact of extended practice on rate of mental rotation. *Journal of Experimental Child Psychology*, *42*(3), 378-391.

Kail, R., & Park, Y. S. (1990). Impact of practice on speed of mental rotation. *Journal of Experimental Child Psychology*, *49*(2), 227-244.

Karwowski, M., & Gralewski, J. (2013). Threshold hypothesis: Fact or artifact?. *Thinking Skills and Creativity*, *8*, 25-33.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children*. John Wiley & Sons, Inc.

Kent, D., & Hedger, K. (1980). Growing tall. *Educational Studies in Mathematics*, *11*(2), 137-179.

Kim, K. H. (2005). Can only intelligent people be creative? A meta-analysis. *Prufrock Journal*, *16*(2-3), 57-66.

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*(5), 745-756.

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamaki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology*, 20, 65 – 76.

Lachance, J. A., & Mazzocco, M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and individual differences, 16*(3), 195-216.

Liben, L. S., & Downs, R. M. (1989). Understanding maps as symbols: The development of map concepts in children. *Advances in child development and behavior*, *22*, 145-201.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development*, 1479-1498.

Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). Top 1 in 10,000: a 10-year follow-up of the profoundly gifted. *Journal of applied Psychology*, *86*(4), 718.

Markey, S. M. (2009). *The Relationship between Visual-Spatial Reasoning Ability and Math and Geometry Problem-Solving*. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.

Mazzocco, M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of dyslexia*, *53*(1), 218-253.

McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children’s arithmetical performance. *Educational and Child Psychology, 20*(3), 93-108.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of experimental child psychology*, *74*(3), 240-260.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental psychology*, *49*(2), 270.

Meyer, M. L., Salimpoor, V. N., Wu, S. S., Geary, D. C., & Menon, V. (2010). Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. *Learning and Individual Differences*, *20*(2), 101-109.

Mix, K. S., Levine, S. C., Cheng, Y., Young, C. Hambrick, Z., Ping, R., & Konstantopoulos S. (2016) The latent structure of spatial skills and mathematics across development: Separate but highly correlated domains. *Journal of Experimental Psychology: General,* 145(9), 1206

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type?. *Personality and Individual Differences*, *50*(8), 1238-1242.

Newcombe, N. S., & Frick, A. (2010). Early education for spatial intelligence: Why, what, and how. *Mind, Brain, and Education*, *4*(3), 102-111.

Noel, M. P., Seron, X., & Trovarelli, F. (2003). Working memory as a predictor of addition skills and addition strategies in children. *Current Psychology of Cognition*, *22*(1), 3-24.

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development*, *22*(2), 165-184.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain and cognition*, *28*(1), 39-58.

Preckel, F., Holling, H., & Wiese, M. (2006). Relationship of intelligence and creativity in gifted and non-gifted students: An investigation of threshold theory. *Personality and individual differences*, *40*(1), 159-170.

Presmeg, N. C. (1986). Visualisation and mathematical giftedness. *Educational studies in mathematics*, *17*(3), 297-311.

Presmeg, N. (2014). Contemplating visualization as an epistemological learning tool in mathematics. *ZDM*, *46*(1), 151-157.

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*(2), 110-122.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of experimental child psychology*, *91*(2), 137-157.

Reuhkala, M. (2001). Mathematical skills in ninth-graders: Relationship with visuo-spatial abilities and working memory. *Educational Psychology*, *21*(4), 387-399.

Robertson, K. F., Smeets, S., Lubinski, D., & Benbow, C. P. (2010). Beyond the threshold hypothesis even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science*, *19*(6), 346-351.

Runco, M. A., & Albert, R. S. (1986). The threshold theory regarding creativity and intelligence: An empirical test with gifted and nongifted children. *Creative Child and Adult Quarterly*, *11*(4), 212-218.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective?. *Psychological bulletin*, *138*(4), 628.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation evidence for multiple representations of numerical quantity. *Psychological science*, *14*(3), 237-250.

Simmons, F. R., Willis, C., & Adams, A. M. (2012). Different components of working memory have different relationships with different mathematical skills. *Journal of experimental child psychology*, *111*(2), 139-155.

Sorby, S. A., & Baartmans, B. J. (2000). The Development and Assessment of a Course for Enhancing the 3-D Spatial Visualization Skills of First Year Engineering Students. *Journal of Engineering Education*, *89*(3), 301-307.

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *100*(2), 343.

Tartre, L. A. (1990). Spatial orientation skill and mathematical problem solving. *Journal for Research in Mathematics Education*, 216-229.

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.

Trbovich, P. L., & LeFevre, J. A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition*, *31*(5), 738-745.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. *Psychology of learning and motivation*, *57*(2), 147-181.

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.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*.

van der Sluis, S., van der Leij, A., & de Jong, P. F. (2005). Working memory in Dutch children with reading-and arithmetic-related LD. *Journal of Learning Disabilities, 38*(3), 207-221.

Van Garderen, D., & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research & Practice, 18*(4), 246-254.

Vasta, R., Knott, J. A., & Gaze, C. E. (1996). Can spatial training erase the gender differences on the water-level task?. *Psychology of Women Quarterly*, *20*(4), 549-567.

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.

Wallace, B., & Hofelich, B. G. (1992). Process generalization and the prediction of performance on mental imagery tasks. *Memory & Cognition*, *20*(6), 695-704.

Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., & Maerlender, A. (2004). Wechsler Intelligence Scale for Children Fourth Edition–Integrated: Technical and interpretative manual. *San Antonio: The Psychological Corporation*.

Young, C. J. & Opfer, J. (2011). Psychophysics of number representation: A unified approach to single and multi-digit magnitude estimation. *Journal of Psychology, 219*, 58-63.

Zimmerman, W., & Cunningham, S. (1991). Editors' introduction: What is mathematical visualization? In W. Zimmerman & S. Cunningham (Eds.), Visualization in teaching and learning mathematics (pp. 1–7). Washington, DC: Mathematical Association of America.