Distribution Agreement

Ye Jin Sung April 10, 2017

Does Mental Rotation Training Improve Arithmetic Competence in Children?

by

Ye Jin Sung

Stella Lourenco
Adviser

Department of Psychology

Stella Lourenco

Adviser

Sherryl Goodman

Committee Member

Robert Jensen

Committee Member

2017

Does Mental Rotation Training Improve Arithmetic Competence in Children?

By

Ye Jin Sung

Stella Lourenco
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2017

Abstract

Does Mental Rotation Training Improve Arithmetic Competence in Children?

By Ye J. Sung

Spatial ability is strongly correlated with mathematic competence and plays a crucial role in children's education (Uttal, Miller, & Newcombe, 2013), specifically in STEM (Science, Technology, Engineering, and Math). However, the causal link between spatial and mathematical reasoning is largely unstudied, limiting the practical implications of the potential relationship. The current study tests whether there is a causal relation between spatial and mathematical abilities by training children on mental rotation or language skills and assessing the impact of training on their math performance. In addition, this study aimed to develop and test the efficacy of a novel 7-day online-at-home training intervention by demonstrating the feasibility of an online procedure for 6- to 7-year-old children. Results suggested that the training groups were successful such that children in the spatial training group showed improvement on mental rotation measures and children in the language training group showed improvement on the language measure. However, we found no evidence of transfer of enhanced spatial performance to mathematical competence. These findings have important theoretical and practical implications and may impact future training studies and real-life applications.

Does Mental Rotation Training Improve Arithmetic Competence in Children?

By

Ye Jin Sung

Stella Lourenco

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2017

Acknowledgements

I would like to express my sincerest gratitude to my advisor Dr. Stella Lourenco for sharing her valuable knowledge and guidance throughout this research project. I would like to extend my gratitude to Chi Cheung as well for her continued support she has provided in completing this project. I also thank my committee members Dr. Sherryl Goodman and Dr. Robert Jensen for their time and insight.

In addition, I would like to express my deepest appreciation to my family and friends for their unwavering support, encouragements, and love. Finally, I acknowledge that this research was supported by a SIRE grant from Emory University.

Table of Contents

Results from the largest cross-national tests, the Programme for International Student Assessment (PISA), which tested 15-year-olds from developed and developing countries, reveals that United States (US) students fall dramatically behind other industrialized nations in mathematical competency (Pew research and American Association for Advancement of Science, 2015). Given these findings, there is increasing concern that the US will not be able to meet the challenges of educating enough citizens to fulfill the demands associated with the STEM (Science, Technology, Engineering, and Math) workforce (Uttal & Cohen, 2012). The lack of strength in STEM-related fields has become such a national priority that, in 2010, President Obama stated: “strengthening STEM education is vital to preparing our students to compete in the 21st century economy and we need to recruit and train math and science teachers to support our nation’s students” (White House Press Release, September 27, 2010). The current method of teaching math is in need of an alternative to current practices and developmental psychologists should be heeding this call by rigorously investigating effective ways to deliver pragmatic interventions.

Previous literature maintains that spatial ability predicts not only the likelihood to pursue but also the likelihood of success in STEM-related careers (Wai, Lubinski, & Benbow, 2009). Other research confirms that adolescents with superior spatial ability earn more STEM-related degrees measured 30 years later (Lubinski, 2010). Though testing spatial ability is not a standard measure for assessing students’ scientific strengths, it has shown to effectively measure an individual’s ability to think scientifically (Shea, Lubinski, & Benbow, 2001). These findings may be explained by the well-established associations between spatial reasoning and mathematical competence (Lubinski, 2010; Wai, Lubinski, & Benbow, 2009; Mix & Cheng, 2012). Indeed, this robust relationship is one of the most widely accepted concepts in the field of spatial cognition (Cheng & Mix, 2012).

Existing evidence suggests that spatial ability, particularly mental rotation, and math share processes especially in the early stages of development (Casey, Nuttall, Pezaris, & Benbow, 1995). Mental rotation is the ability to manipulate and rotate images in the mind, a skill essential to solving a number of math problems like algebra, geometry, and even general calculation. Many studies point to male's superior mental rotations skills detected as early as 3-to-5 month old infants for the observable gender differences in math achievement (Moore & Johnson, 2008; Quinn & Liben, 2008). This is because spatial reasoning, including mental rotation, measured as early as 6 months of age is found to predict the development of math concepts measured at 4 years (Lauer & Lourenco, 2016).

The early observed differences in spatial ability, like mental rotation, seems to have a lasting impact on student's math aptitude. In addition, there exists a strong relationship between early number competence and later mathematic achievement (Jordan, Kaplan, Ramineni, & Locuniak, 2009). Therefore, nurturing mental rotation abilities, especially during developmental periods may be critical for closing the achievement gaps in math. A 5-year longitudinal study investigated whether testing for spatial ability can reveal math and science aptitude by assessing individual spatial abilities and comparing them with Scholastic Aptitude Test (SAT) scores for math and verbal in 7th to 10th grade students (Webb, Lubinski, & Benbow, 2007). Results showed a positive relationship between spatial ability and SAT-Math scores even when accounting for SAT-Verbal scores. Lubinski (2010) also reported that mental rotation tasks are positively correlated with math achievement among children from Kindergarten to 12th grade. Again, illuminating that the effects of spatial ability remain stable across the life span.

Given this information, Webb, Lubinski, and Benbow (2007) suggested that we will be able to enhance the currently hidden talents of potential mathematicians and scientists by fostering spatial ability. Tests of spatial ability have the potential to identify talent that current standardized math testing may not capture; thus, incorporating spatial training will be helpful for math education by allowing the current educational system to uncover the untapped scientific talent that the US currently lacks.

It has been repeatedly found that spatial abilities are related to mathematical ability, beginning early in development, but it is largely unclear what constitutes these developmental relationships. The existing literature offers us several possible explanations. One prominent rationalization is that numbers are inherently spatial; spatial reasoning and math competence have such a strong association because numbers are deeply fixed in space in the mind. The mental number line theory, broadly explains that numerical quantities are represented in spatial format in the brain; numbers and their representative magnitudes are automatically linked to a mental number line primarily aligned from left to right in the Western cultures (Dehaene, Bossini, & Giraux, 1993; Fias & Fischer, 2005; Moyer & Landauer, 1967). Numbers representing small magnitudes will be spatially conceptualized towards the left side of the number line while numbers representing larger magnitudes will be spatially conceptualized to the right side of the cognitive number line.

This phenomenon is well exemplified by Dehaene el al. (1990) where they presented a reference number followed by a number stimuli. Then they asked participants to indicate if the stimuli were smaller or larger than the reference number by pressing either the left key or the right key. As a result, participants who were asked to press the left key for the smaller number and the right key for the larger number had significantly faster response times. Thus, confirming that the mental number line is directed from the left to right in magnitude. This association was named the SNARC (Spatial Numerical Association of Response Codes) effect with numerous studies that follow to verify it (Dehaene et al., 1993; Fias, Brysbaert, Geypens, & D'Ydewalle, 1996; Nuerk, Wood, & Willmes, 2005). Neural data also confirms the association between space and numbers. Brain imaging scans reveal similar areas of the parietal lobe are activated when processing abstract number representations and spatial information (Hubbard, Piazza, Pinel, & Dehaene, 2005; Dehaene, Piazza, Pinel, & Cohen, 2010). In addition, there exist evidence that spatial processes support math reasoning by grounding abstract concepts in concrete spatial format or by providing mental models to solve specific math problems (Cheng & Mix, 2012; Lauer & Lourenco, 2016).

Spatial reasoning skills are highly valuable even outside the realm of STEM abilities; spatial ability promotes functioning in areas that may not seem spatial (Newcombe & Frick, 2010; Huttenlocher & Lourenco, 2007). Therefore, a student not interested in pursuing a STEM related education would also benefit from spatial skills training. Research indicates that spatial skills are applied in our daily life: trying to navigate our ways through unfamiliar road, trying to fit things together most effectively in a small space, changing lanes among high-speed traffic, reasoning about location, and even in assembling objects from separate parts.

It is imperative to understand that spatial ability is malleable and can be improved by training (Heil, Rösler, Link, & Bajric, 1998; Ehrlich, Levine, & Goldin-Meadow, 2006; Newcombe & Frick, 2010; Uttal et al., 2012). A meta-analysis in particular aimed to offer a comprehensive look by analyzing 217 studies that specifically investigated spatial skill training (Uttal et al., 2012). They confirmed that spatial reasoning can be improved through training with the average effect size of 0.47 (Hedges's g). These results were strong as they held to be true even when accounting for different settings, method of training, length of training, and discrepancies in gap days between last training and post-test administration for each study. Moreover, this meta-analysis also found that spatial training transferred beyond the training itself and to other non-targeted spatial skills, which suggests the possibility that spatial training may have an impact on numerical abilities as well.

The relationship between space, math, and the malleability of spatial cognition is clearly robust enough to be of interest to current educators and researchers to implement in classrooms. Though there is a plethora of research showing the association between spatial ability and math reasoning, most of these studies provide correlational evidence. As it is well known that correlation does not imply causation, to establish the causal relation, training research is needed to demonstrate that improvement in spatial abilities will lead to improvement in mathematics. To the best of our knowledge, there are two published training studies, but they have produced inconsistent findings. Whereas one study found that mental rotation training improved arithmetic performance (Cheng & Mix, 2012) the other did not (Hawes, Moss, Caswell, & Poliszczuk, 2015). Mental rotation is a hallmark of spatial reasoning that involves the ability to manipulate, specifically rotate, figures in one's mind. Given the theoretical importance of this issue and the discrepancies in the current literature, it is important to examine whether mental rotation training indeed improves arithmetic ability, and if so, what are the critical elements that make such kind of training effective?

In the first study, Cheng and Mix (2012) tested 6- to 8-year-old children to investigate whether mental rotation training would lead to improvement in mathematics. A pre-test was followed by a training session of 40 minutes. Then, immediately after training, they administered the post-test. The researchers found that a single session of mental rotation training in the lab led to significant improvement in arithmetic performance. However, it is important to note that out of the three subsets of the math test (e.g. missing term, multi-digit, and number fact), Cheng and Mix found significant improvements only on the missing-term problems (e.g., 4+__= 12). In addition, they did not find improvement in mental rotation performance broadly.

One notable feature of this study was that immediate corrective feedback was included in the training; participants had the opportunity to check their work by physically putting together cardboard shapes identical to the corresponding training stimuli. This suggests that direct instruction and feedback may be a critical element that contributes to transfer from spatial training to improved arithmetic performance. Nevertheless, the question remains as to why they did not find more general spatial improvement since these tests would have been more similar to the training intervention than the missing-term test. It is strange to see effects of far transfer when near transfer didn’t occur and these concerns bring to question the effectiveness of training as well as priming effects.

Hawes, Moss, Caswell, and Poliszczuk (2015) conducted a similar study on 6- to 8-year-olds to replicate and extend the above study. They tested whether a computerized mental rotation training could enhance spatial abilities and how much spatial training would transfer to children’s calculation performance before and after 6-weeks of training. The training was administered in an elementary school setting with teachers monitoring the progress of their students. Critically, children received no feedback throughout the training. As a result, Hawes et al. found no observable improvement in calculation and missing term problems even though they found improvement in general spatial abilities. More specifically, they assessed and found immediate and near transfer effects of spatial ability. In contrast to Cheng and Mix’s findings, improvements in one specific spatial skill transferred to improvements in other spatial skills as well. Despite meeting the both criteria of successful improvement in trained items and transfer to general spatial skills, Hawes et al. did not find any evidence that suggest spatial training yields improvement in mathematical skills.

Given the conflicting findings, the value of utilizing spatial training to directly enhance mathematic competency requires more attention. We hope to combine the strengths of the two studies in order to answer some questions that the inconsistent results raised. Cheng and Mix only implemented a 40-minute training and administered the post-test immediately afterwards. Hawes et al. employed 6-week training with the post-test 3-6 days after training. Therefore, we found a feasible middle ground of administering a pre-test, 7-day online mental rotation intervention, and a post-test approximately the day after the last training session.

For the purposes of this study, we had aimed to include only the spatial training group due to the magnitude of this study. However, a control group is important in ruling out practice effects; thus, we strived to include as many children as possible in the control group that received the language training. This group is preliminary and will require further recruitment. Although training was delivered via a website, corrective feedback was immediately provided upon participants' responses. This was incorporated purposefully to test whether direct instruction is the critical element that makes spatial training effective. In addition, instructional videos were employed to ensure learning. Because, Cheng and Mix (2012) found that mental rotation training led to improvement in missing-term problems specifically, we also tested children on this type of math problems. Importantly, we also included measures of near transfer. That is, children were tested on a spatial assessment they were not specifically trained for, Picture Rotation (Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011), to assess the extent to which our training improved their general mental rotation ability. Additionally, we also included assessments of working memory as a general intelligence measure and Approximate Number System (ANS) to test non-symbolic math skills which requires a combination of numerical and spatial skills. These measures were not included in the past two studies but we aimed to test if mental rotation training could also show transfer to ANS acuity because it is closely related to both math and space. Likewise, the general intelligence measure was included to see whether it had any influence on the effectiveness of the training.

Method

Participants

Thirty, typical-developing children, (M_{age}= 6.66, SD_{age}= 0.35) were recruited from the greater Atlanta area. The experiment group, receiving spatial training, consisted of 22 participants with one drop out. The control group, receiving reading training, consisted of 7 participants. Of the 30 participants, 29 were included in the final analysis where 13 participants were female and 16 participants were male. Although demographic information was not analyzed systematically, the majority of participants were part of families that are high in socio-economic status.

As this study requires a relatively large commitment (6 at home training days and 2 visits to the lab), we provided monetary compensation of $40 to parents upon completion of the study and all children received gifts for participating. Informed consent was obtained on behalf of each child by a parent or legal guardian. Experimental procedures were approved by the local ethics committee.

Materials and Procedure

All participants were assessed with the same battery of tasks in the pre-and post-test sessions. Each participant came into the lab for the first session, pre-test, to assess the baseline for all tasks. Then they returned to the lab approximately a week later for the post-test after the completion of all 7 days of the online training. Therefore, the participants would return the day after their last day of training, leaving no gap days in between the last day of training and the 2nd test session. The order of individual tasks was randomized within the test blocks. Half of the participants completed the math assessment block first and the rest completed the spatial assessment block first. The last testing block was always the general intelligence block. Upon completion of pre-test assessments, participants completed the first training session under the experimenter's supervision. This procedure ensured that all participants fully understand how to complete the training at home. The remaining training of 6 additional days took place at home using our online platform under the supervision of their parents.

Assessment of Arithmetic Competence

The Woodcock Johnson [WJ] Test of Calculation (Woodcock, McGrew, & Mather, 2001), a typical test for children in this age, was used as one of the two assessments of arithmetic performance. Participants were provided with a worksheet of standardized calculation problems with no time constraint. The test started with single digit addition and subtraction and became progressively difficult. The participants were asked to start from top and complete all the problems they knew how to solve and stopped when they answered 6 consecutive questions incorrectly.

A test of missing-term problems was created to mirror the test used in the study of Cheng and Mix (2012). We presented a total of 12 missing-term problems on a single sheet of paper with the first one as the practice question (See Figure 11). Problems were carefully constructed to include the same number of addition and subtraction problems. We also varied the position of the missing term (e.g., __+ 5 = 7 and 2 + __ = 7) so that there was an equal number of each type of questions. The first question was explained and solved together by the participant and the experimenter. Further clarification and correction was provided if needed on this problem. After the participants understood the task, participants were given as long as they needed to complete the remaining 11 questions on their own and no feedback was given. The practice problem was not scored and the rest of the 11 questions were scored as either correct or incorrect. Correct answer equaled one point and incorrect answer equaled zero points and the total score was the sum of these points.

We also included an assessment of the Approximate Number System (ANS) designed following previous studies (Temple & Posner, 1998). The test measures the individual's cognitive sensitivity to estimate and compare two sets of numerosity without exact counting. This measure is different from the above math measures because it is completely non-symbolic yet still involve numerosity. Two sets of circles, unique for every test item, representing different ratios of magnitude were presented for 2 seconds on the computer screen. Then, the participants were asked to estimate the magnitude of each group and pick which set of circles are greater. Again, the total score was the sum of the correct answer being one point and the incorrect answer being zero points.

Assessment of Spatial Ability

Children's Mental Transformation Task (CMTT) was used to measure children's mental rotation performance. The test was created by Levine et al. (1999) and was administered by both Cheng and Mix (2012) and Hawes et al. (2015). Each item includes a target shape and 4 choice arrays (see Figure 1). The total set of 32 items included four types of rotations: direct translation, diagonal translation, direct rotation, and diagonal rotation (See Figure 2). The original task created by Levine et al. (1999) included 32 questions. We randomly picked 4 items of each category, total of 16, to include in our pre- and post-tests. Participants were told that the two pieces on top come together to make one of the shapes on the bottom. They chose their answer on a touch screen computer with no time constraint and each correct test question was awarded one point.

The Picture Rotation Test (see Figure 3) is a mental rotation test where children are presented with one target animal on the left side with two of the same animals rotated in different positions and two mirror images of the same animals on the right (Neuburger et al., 2011). Participants were asked to identify and circle which two of the four choice animals were rotated versions of the target animal while avoiding the mirrored images of the target item. The experimenter explained the task with two examples, and then the participant was allowed two practice problems to ensure that they understand the task. Participants were corrected on the practice problems and any questions were answered to clarify the task. There were a total of 16 test questions and participants were given 2 minutes to complete as many as they could. The participant had to circle both rotated animals in order to score one point.

Additional Measures

We also included measures of general intelligence. Specifically, we assessed working memory as it is the capacity to remember and recall relevant information. Working memory was measured with the Woodcock Johnson [WJ] Auditory Working Memory measure (Woodcock, McGrew, & Mather, 2001), which tests individual’s memory span. The task requires that participants hear sets of words and numbers in mixed order and reiterate them back, providing the words first in the same order and followed by the numbers in the same order. The second measure was the Woodcock Johnson [WJ] Word Identification test (Woodcock, McGrew, & Mather, 2001), which was used to assess language cognition. Participants were asked to read sets of words that grew increasingly difficult. The participants must pronounce the words correctly to score a point. The items were administered until the participant could not pronounce any of the listed words in one page.

Online Training

Participants were either assigned to an experimental or control group; the process was non-randomized as all participants were placed in the experimental group primarily and the control group was formed after. This was to ensure that we had enough participants in the experimental group. The experimental group received mental rotation training and the control group received a language training that involved neither spatial nor mathematical abilities. The control group was created to confirm that any improvements observed with this group of children were specific to mental rotation training and not simply the result of playing an online game per se. Parents received daily reminders about the training via email or text message.

Each mental rotation training session included an instructional video (approximately 30 seconds), which was followed by eight practice items from the CMTT. In each trial, two pieces would be shown above the four different shapes (see figure 9). The first training session only contained translation because we determined that it was the easiest mental rotation out of the four types – it required the least amount of effort to put them together. All subsequent training sessions consisted of two items of each type. Corrective feedback was provided after participant made their choice.

It is important to note that some of the CMTT training items were not the original target items of the CMTT. This procedural decision was made because excluding the 16 items included in our pre-test and post-test left us with only 16 items. However, for the purpose of our training, 7 days with 8 CMTT sets, we required more than 16 items. Each question had 4 choices with one target shape and three distractors. Therefore, we created additional training stimuli by splitting each of the three distractor shapes to recreate them as target shapes. As a result, we had a total of 56 questions even after excluding the 16 questions we reserved for the pre-and-post-test questions.

The control group's language training was the same spatial training except the task was to identify the target word out of four choice of words (see figure 10). The pronunciation of the target word was given and participants had to choose the correct answer. It also included instructional videos, direct corrective feedback, and 8 training items per day.

Results

Training Completion

Throughout the experiment, the progress of training completed by participants were thoroughly tracked to make sure they received the appropriate amount of training sessions. Almost all participants finished training properly with minor deviations. Out of the 23 total participants in the spatial intervention group, 20 participants completed all 7 training sessions successfully and 2 participants completed 5 training sessions. One participant dropped out of the study during training. Because analyses revealed that participants who completed 7 and 5 training sessions did not appear to differ from each other, we treated all children within this group similarly in the analyses reported below. In the language training group, all 7 participants successfully completed all 7 training sessions.

Pre-test Comparisons

As a first step in assessing between-condition differences (experimental vs. control groups), we tested whether children assigned to the experimental and control groups differed from one another at the pre-test period. Using independent-samples t-tests, we found no significant differences at the pre-test period for any of the measures (i.e., CMTT, PRT, Missing Term Problems, Calculation, and Working Memory). The ANS test, however, showed a significant difference between spatial (M= 15.77 SD= 1.85) and language training groups (M = 16.86, SD = 0.69), $t(25.99) = 2.29$, $p = 0.03$ (corrected for unequal variances). The participants were high performers in general; However, children in the reading group in particular performed higher on all the measures except for the Word Identification Test.

Did spatial training work?

First, we conducted a within-group paired samples t-test to ensure that the online training intervention had achieved its purpose – namely, to enhance the targeted spatial ability. In the current study, we tested specifically for whether spatial training would increase mental rotation ability and whether language training would increase language ability.

When pre- and post-test scores of the CMTT were compared in the spatial training group ($M_{difference\ score}$= 2.86 $SD_{difference\ score}$= 2.88), the results indicated that children in this group showed significant improvement in mental rotation measured by their CMTT scores $t(21) = 4.66$, $p < 0.001$, d= 0.99, whereas the language training group ($M_{difference\ score}$= 0.71 $SD_{difference\ score}$ = 2.75) did not $t(6) = 0.69$, $p = 0.52$ (see Figure 6). In a second analysis, we compared pre- and post-test scores on the CMTT between spatial training and control groups. More specifically, we tested whether the difference between pre- and post-test in the spatial training group represented a bigger improvement than in the control group. This analysis revealed a marginally significantly effect, *t*(27) = 1.74 , p = 0.09, d=0.68. Taken together, these findings suggest that only the spatial training group showed significant improvement between pre- and post-test, though this difference was not large enough to meet statistical significance when comparing it to children's performance in the control group.

The picture rotation (PRT) task, our second measure of spatial ability, would capture whether or not the CMTT training offered a more general enhancement of spatial ability. A paired samples t-Test on the picture rotation (PRT) revealed that the spatial training group ($M_{difference\ score}=$ 4.82 $SD_{difference\ score}$= 2.61) improved significantly, *t*(21) = 8.65, $p <$ 0.001, d= 1.84. However, the language training group ($M_{difference\ score}=$ 2.14 $SD_{difference\ score}$= 1.77) also improved significantly, *t*(5) = 2.74 , p = 0.02, d=1.21. Even though both spatial and language group improved for the picture rotation (PRT) task, the spatial training group's improvement was considerably higher (see Figure 7).

In the second analysis, I took the difference score of the pre-and-post-test for Picture Rotation (PRT) to compare those scores between spatial and control groups. This is to see if the improvement in spatial group was significantly greater than the improvement found in the reading group. Remarkably, this effect was found to be true even when increased scores are compared between the two groups; *t*(27) = 2.52, p = 0.018, d= 1.09. Taken together, these findings suggest that although both training groups showed improvement from pre- to post-test, there was greater improvement in the spatial training group, suggesting a benefit of the spatial training group as was found for the other measure of mental rotation (CMTT). Thus, across two measures of mental rotation, results indicated that the spatial training group improved in performance.

Was working memory responsible?

Now that the results indicate success of spatial training, we aimed to test whether the increase in spatial ability scores was specific to improved mental rotation ability by controlling for working memory scores. One possibility is that children showed improvement in the spatial training group, not because their mental rotation ability was enhanced as a result of training, but because the training improved more general processing abilities such as working memory. An analysis of covariance (ANCOVA) with the dependent variable as the difference scores of CMTT and working memory as the covariate was conducted. The analysis found no significant relation between the CMTT improvement and Working Memory improvement, $F(1, 26) = .729$, $p = .401$. The effect of training was not significant, $F(1,26) = 2.15$, $p = .154$. However, it should be noted that the effect size for training was medium to large (η_p^2= .07), while the effect size of working memory improvement was small to medium (η_p^2= .03). Thus, it is unlikely that working memory improvement is responsible for the improvement in CMTT. Rather, adding working memory as a covariate further decreased the statistical power of the analysis, leading to the insignificant effect of training. A similar ANCOVA analysis was conducted on the difference scores of PRT, with working memory improvement as covariate. The analysis also found that the amount of working memory improvement was marginally related to PRT improvement, $F(1, 26) = 4.17$, $p = .051$. However, most importantly, the effect of training remained significant after working memory improvement was controlled, $F(1, 26) = 9.46$, $p = .005$. Thus, the improvement in Picture Rotation Test (PRT) cannot be attributed to improvement in working memory.

Did the Language intervention work?

Even though the language intervention was for the control group, we also analyzed the word identification test if the language training worked. First, we conducted a within-group paired samples t-test. When the pre-and-post test scores of the word identification test were compared within the language group ($M_{difference\ score}$= 2.57, $SD_{difference\ score}$= 2.23), results indicated significant improvement $t(6) = -3.06$, $p = 0.022$, $d = 1.16$ in contrast to the spatial group ($M_{difference\ score}$< 0.00001, $SD_{difference\ score}$= 3.06), $t(21) < 0.001$, $p = 1.00$ (see Figure 8). In the next analysis, the pre-and-post-test scores for the Word Identification test were compared between the spatial group and the language group. The analysis showed significant results $t(27) = -2.05$, $p = 0.050$, $d = 0.89$, indicating that the improvement of the language group was significantly greater than the improvement of the spatial group in the Word Identification scores. Therefore, the data suggest that the language group's training achieved its goals of improving language skills.

Was working memory responsible?

Similar to the spatial training we also tested whether the improvement in word identification is caused by improvement in WM. We analyzed the difference scores of Word Identification Scores conducted using ANCOVA, with intervention as the between subject factor improvement in working memory as covariate. The analysis indicated a non-significant effect between the Word Identification Test improvement and Working Memory improvement, $F(1, 26) = 2.42$, $p = 0.13$ and the effect of training was also non-significant $F(1, 26) = 2.76$, $p = .10$. Even though results seem to suggest working memory explains the training effect, the intervention had a medium to large effect size (η_p^2 = 0.10), and the results need to be interpreted with caution.

Doe spatial training improve competence?

Missing Term Problems

We tested these specific set of missing term problems, as Cheng and Mix (2012) emphasized that they found a particularly significant increase in missing term problems. Again, the paired samples t-test was conducted to determine whether groups improved from their pre- to post-test scores. The spatial group ($M_{difference\ score}$= 0.41, $SD_{difference\ score}$=1.22) did not improve significantly, $t(21) = 1.57$, $p = 0.13$ and neither did the reading group ($M_{difference\ score}$= -0.06, $SD_{difference\ score}$=1.27), $t(6) = 0.89$, $p = 0.41$ (see Figure 5). An independent sample t-test was also conducted to detect significant differences in score increase when comparing the spatial group and the language group. The dependent variable was set as the difference in missing term scores from pre-to-post-test. However, we found no significant increase in scores across groups, $t(27) = 1.55$, $p = 0.13$.

Calculation

The Woodcock [WJ] Calculation test was also analyzed. The paired samples t-test showed no significant improvement in pre- to post-test scores for neither spatial group ($M_{difference\ score}$= 3.00, $SD_{difference\ score}$ = 11.59), $t(21) = 1.21$, $p = 0.24$, nor for the reading group ($M_{difference\ score}$= -2.71, $SD_{difference\ score}$= 6.99), $t(6) = 1.03$, $p = 0.34$ (see Figure 4). Another analysis of independent sample t-test was carried out with the difference scores of the calculation scores as the dependent variable. There was no significant effect, $t(27) = 1.226$, $p = 0.23$. These results indicate that there was no significant increase in mathematical abilities when the difference scores of the spatial group were compared to the language group. The data suggests that there was no evidence of transfer.

Discussion

The current study aimed to build on existing literature and further explore the causal links between spatial ability and mathematical competence by training children in spatial ability. In addition, we hoped to develop a novel training technique. These two aims go hand in hand because training studies are necessary to investigate direct causal relationships. Despite the importance of causal evidence, there may exist such few training studies because training studies are intensive and notoriously difficult to conduct. However, our novel online-at-home approach to training may offer a more convenient way to implement intervention.

The data indicate that both spatial training and language training were successful in enhancing their respective targeted abilities. Although the CMTT showed marginal significance in our preliminary results, this trend is very promising given that the sample sizes were small and unbalanced. We are currently underpowered but may be able to detect a significant effect once more participants are recruited. The effects were seen not only in the directly trained test of CMTT, but also a non-trained spatial measure of PRT, which indicates that the training genuinely improved spatial ability.

This new method of training works and the majority of our participants finished training without any difficulties. Parents repeatedly reported that children were highly engaged and enthusiastic about their daily training regardless of their starting mathematic or spatial competence. Some participants even expressed their desire to continue playing the training game after the experiment was over. The success of our training demonstrates that a training intervention is attainable at home under parental supervision. By handing over the supervision component to the parents and reducing the burden of coming in to the lab makes training studies more achievable. These results are very promising and provide a new method in conducting future training studies.

In regards to mathematical competence, our preliminary data detected no evidence of math transfer in either missing term problems or a standardized calculation test. Despite the successful intervention, our data contributes to the literature that supports training spatial ability and succeeding does not necessarily enhance arithmetic competence. However, this is only one more study on top of the two existing studies and more research is needed to confirm these findings.

Given the abundant correlational studies confirming the robust relationship between space and math, this finding may seem counterintuitive. We infer that when training does work, the trained ability is very specific. Even if a specific spatial training such as mental rotation achieves near-transfer of also yielding benefits in general spatial cognition, perhaps the effects do not reach the relatively further skills of general math. A large meta-analysis of working memory training revealed similar findings – training on working memory worked but no transfer occurred to related skills or improved general intelligence (Melby-Lervåg, Redick, & Hulme, 2016). Working memory is widely believed to contribute to high functioning in many areas of academic and everyday functioning. As general as working memory is, the meta-analysis still failed to find lasting far-transfer effects. Even when they detected a significant effect, the transfer was only temporary; this may explain how Cheng and Mix (2012) found transfer to math – by administering the post-test immediately after the post-test.

Though there are not enough studies to draw a firm conclusion, spatial training may not be very different from the above findings as spatial ability is much more specific than working memory. Furthermore, it may be that mental rotation does train children in spatial ability but may not precisely tap into the cognitive abilities that math and space share. Perhaps we have to figure out exactly what kinds of functions are shared between math and space to specifically target and train those abilities.

However, our data seems to discourage the effects of spatial training in nurturing arithmetic competence. This has important implications because many institutions are investing millions of dollars on spatial training implementation in the hopes of raising mathematic competency in American students. Also, the National Council of Teachers of Mathematics (NCTM) advocates that preK-8th grade math education should largely focus on spatial training (Schwartz, 2017). These monetary and timely investments are based on only the correlational studies that prove spatial ability is related to math. It is critical to understand that correlation does not imply causation and the existing literature, including the current study, does not provide enough causal evidence required to motivate these large investments. Schwartz (2017) states that the teachers started labelling spatial training as math and started to teach spatial reasoning exclusively. We need to recognize the potential dangers of jumping to conclusions. Without sufficient empirical data to favor spatial training enhancing math aptitude, children may be neglected of a proper math education if spatial training start to replace traditional math learning. Improving mathematical education is urgent, but additional investigation of the causal relationship is absolutely essential before we move towards spatial training implementation in classrooms.

Limitations

As the findings of the current study is preliminary and it is to continue, we had some limitations. We had primarily aimed to test only the spatial group due to the time constraints and the intense nature of a training study, but included a control group to genuinely access the effectiveness of our training and transfer to math. Therefore, we had a non-randomized assignment where the first participants were placed in the spatial group until it was filled and then the remaining participants were placed in the language group. The time constraints also explain why the sample size is small and the number of participants in the spatial group and the reading group is unmatched. With a small, unmatched sample, our data analysis was underpowered and it may not have been the most effective comparison. Increasing and matching the sample size is the ultimate goal as we continue to recruit and test participants.

Another concern is the self-selecting bias. As this study is time consuming and requires two visits to our lab, we suspect that parents who agree to participate in our study are those who have great interest in the child's education. Therefore, our sample for both spatial and language group consisted of very high performing children in general. Particularly, the language group contained a cluster of unusually high performers, which raised the concern whether or not this was an effective control group.

Future Directions

The current study implemented an online-at-home training that effectively enhanced mental rotation but found no evidence of transfer to mathematic skills. We will continue to recruit more participants in the efforts to increase the sample size as well as balance the number of spatial group and the language group to address the issue of being under power.

Our intervention technique shows promising results and suggests that such training can maximize benefits by reducing the burden for both experimenter and participants while still producing desired outcomes. Therefore, more studies should explore ways to further develop the efficacy of this type of training to help facilitate more training studies. Furthermore, even though our data seems to discourage the effects of spatial training, the efforts to find innovative ways to teach math should persist. If not mental rotation, it may be another aspect of spatial ability, that will lead to an increase in mathematical ability.

References

Bonny, J. W., & Lourenco, S. F. (2014). Achievement : Evidence from the Preschool Years, *114*(3), 375–388. http://doi.org/10.1016/j.jecp.2012.09.015.The

Cheng, Y. L., & Mix, K. S. (2012). Spatial Training Improves Children's Mathematics Ability. *Journal of Cognition and Development*, *15*(May 2015), 120919075341007. http://doi.org/10.1080/15248372.2012.725186

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*. http://doi.org/10.1037/0096-3445.122.3.371

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology. Human Perception and Performance*, *16*(3), 626–641. http://doi.org/10.1037/0096-1523.16.3.626

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2010). Cognitive Neuropsychology THREE PARIETAL CIRCUITS FOR NUMBER PROCESSING T HREE PARIETAL CIRCUITS FOR NUMBER. *Psychology*, *3294*(April), 37–41. http://doi.org/10.1080/02643290244000239

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology*, *42*(6), 1259–68. http://doi.org/10.1037/0012-1649.42.6.1259

Fias, W., Brysbaert, M., Geypens, F., & D'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition*. http://doi.org/10.1080/135467996387552

Fias, W., & Fischer, M. (2005). Spatial representation of number. *Handbook of Mathematical Cognition*, (JANUARY 2005), 43–54.

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), 1–9. http://doi.org/10.1016/j.tine.2015.05.001

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. http://doi.org/10.1007/s004260050016

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews. Neuroscience*, *6*(6), 435–48. http://doi.org/10.1038/nrn1684

Huttenlocher, J., & Lourenco, S. F. (2007). Huttenlocher, J. & Lourenco, S.F. (2007). In J. Plumert & J. Spencers (Eds.), The Emerging Spatial Mind.pdf. Plumert & J. Spencer (Eds.), The Emerging Spatial Mind. http://doi.org/285-290

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. (2009). NIH Public Access, *45*(3), 850–867. http://doi.org/10.1037/a0014939.Early

Lauer, J. E., & Lourenco, S. F. (2016). Spatial Processing in Infancy Predicts Both Spatial and Mathematical Aptitude in Childhood. *Psychological Science*. http://doi.org/10.1177/0956797616655977

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early Sex Differences in Spatial Skill. *Developmental Psychology*, *35*(4), 940–949. http://doi.org/10.1037/0012-1649.35.4.940

Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences*, *25*(410), 126–133. http://doi.org/10.1016/j.lindif.2013.02.001

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, *49*(4), 344–351. http://doi.org/10.1016/j.paid.2010.03.022

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS ONE*, *6*(9), 1–9. http://doi.org/10.1371/journal.pone.0023749

Melby-Lervåg, M., Redick, T., & Hulme, C. (2016). Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of "Far Transfer": Evidence from a Meta-Analytic Review. *Perspectives on Psychological Science*, *11*(4), 512–534. http://doi.org/10.3837/tiis.0000.00.000

Mix, K. S., & Cheng, Y. L. (2012). *The Relation Between Space and Math. Developmental and Educational Implications*. *Advances in Child Development and Behavior* (Vol. 42). Elsevier Inc. http://doi.org/10.1016/B978-0-12-394388-0.00006-X

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

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, *8*(2), 228–246. http://doi.org/10.1016/0010-0285(76)90025-6

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520. http://doi.org/10.1038/2151519a0

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. http://doi.org/10.1016/j.paid.2011.02.017

Newcombe, N. S., & Frick, A. (2010). Early Education for Spatial Intelligence : Why , What , and How, *4*(3), 102–111.

Nuerk, H. C., Wood, G., & Willmes, K. (2005). The universal SNARC effect: The association between number magnitude and space is amodal. *Experimental Psychology*, *52*(3), 187–194. http://doi.org/10.1027/1618-3169.52.3.187

Pew research and American Association for Advancement of Science. (2015). public and d Scientists' View ws on Science and Society, *1–93*, 1–93.

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

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

Spelke, E. S. (2005). Sex Differences in Intrinsic Aptitude for Mathematics and Science?: A Critical Review. *American Psychologist*, *60*(9), 950–958. http://doi.org/10.1037/0003-066X.60.9.950

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7836–41. http://doi.org/10.1073/pnas.95.13.7836

The White House, Office of the Press Secretary. (2010, September 27). *President Obama Announces Goal of Recruiting 10,000 STEM Teachers Over the Next Two Years* [Press Release]. *The White House.*

Uttal, D. H., & Cohen, C. A. (2012). *Spatial Thinking and STEM Education. When, Why, and How? Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 57). Elsevier Inc. http://doi.org/10.1016/B978-0-12-394293-7.00004-2

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2012). The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. *Psychological Bulletin*, *139*(2), 352–402. http://doi.org/10.1037/a0028446

Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and Enhancing Spatial Thinking: Links to Achievement in Science, Technology, Engineering, and Mathematics? *Current Directions in Psychological Science*, *22*(5), 367–373. http://doi.org/10.1177/0963721413484756

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

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, *99*(2), 397–420. http://doi.org/10.1037/0022-0663.99.2.397

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III.* Itasca, IL: Riverside.

Appendix A

Figure 1. An example of the CMTT mental rotation test and training stimuli used in the experiment (Levine, Huttenlocher, Taylor, & Langrock, 1999). 16 of these items were administered to measure spatial ability. The participants were instructed to choose which of the four shapes on the bottom the two pieces on top makes.

Appendix B

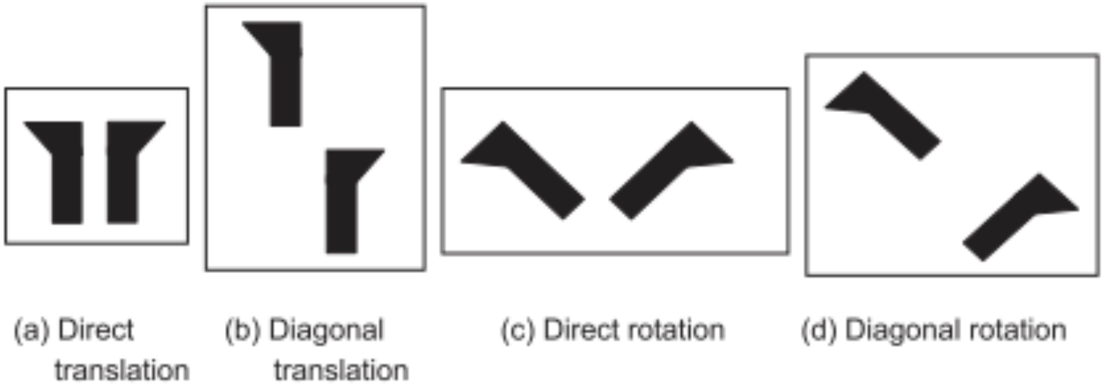

Figure 2. Four types of CMTT test items displayed above the four choice arrays (Levine et al., 1999).

Appendix C

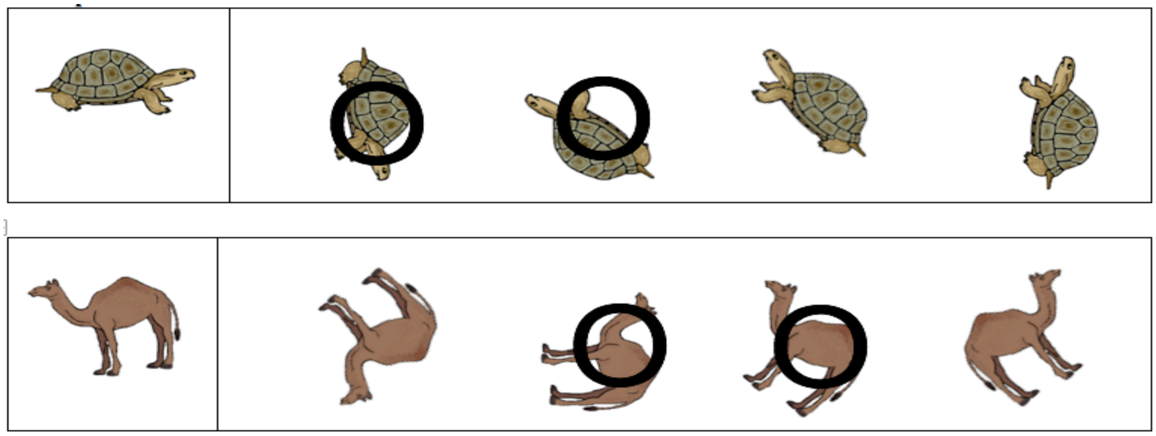

Figure 3. Displays the Picture Rotation Test (PRT) (Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011) administered to measure near transfer of spatial ability. Above picture is the example shown to the participants prior to the scored items. Participants are instructed to discriminate against the mirror images and circle the image that has been rotated.

Appendix D

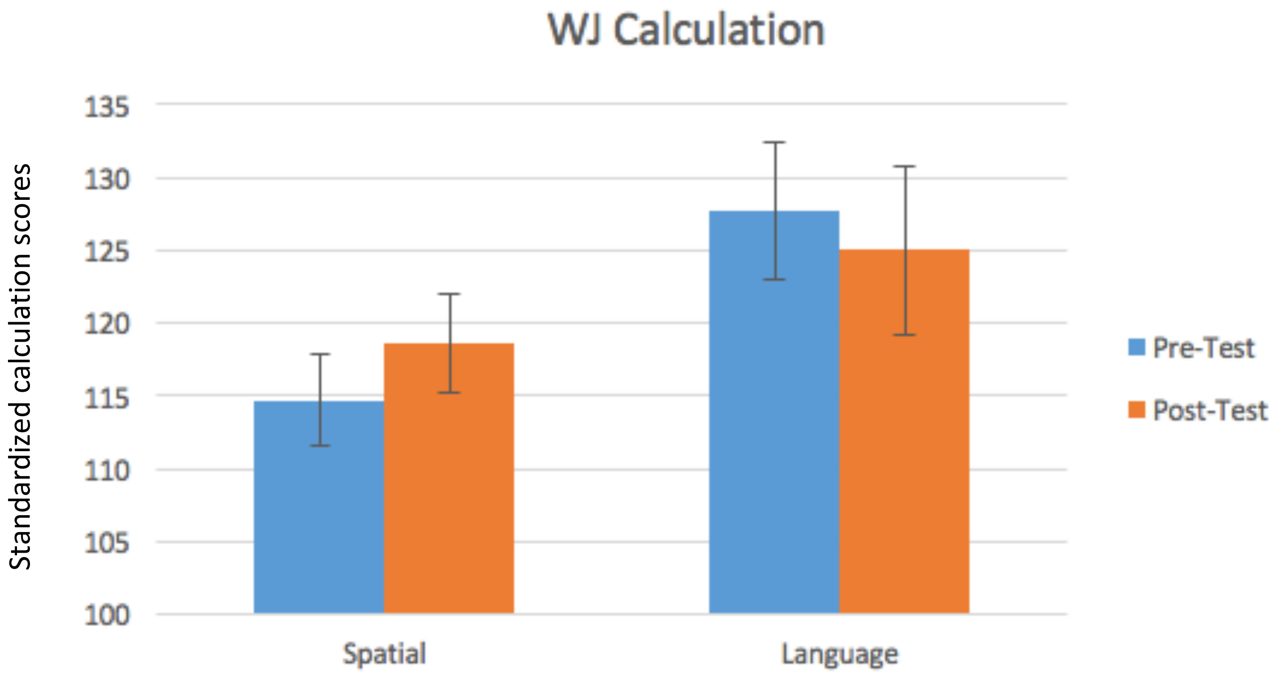

Figure 4. The pre-test and post-test of the Woodcock Johnson [WJ] Calculation scores for both Spatial group and Language group. Error bars represent standard error of the mean.

Appendix F

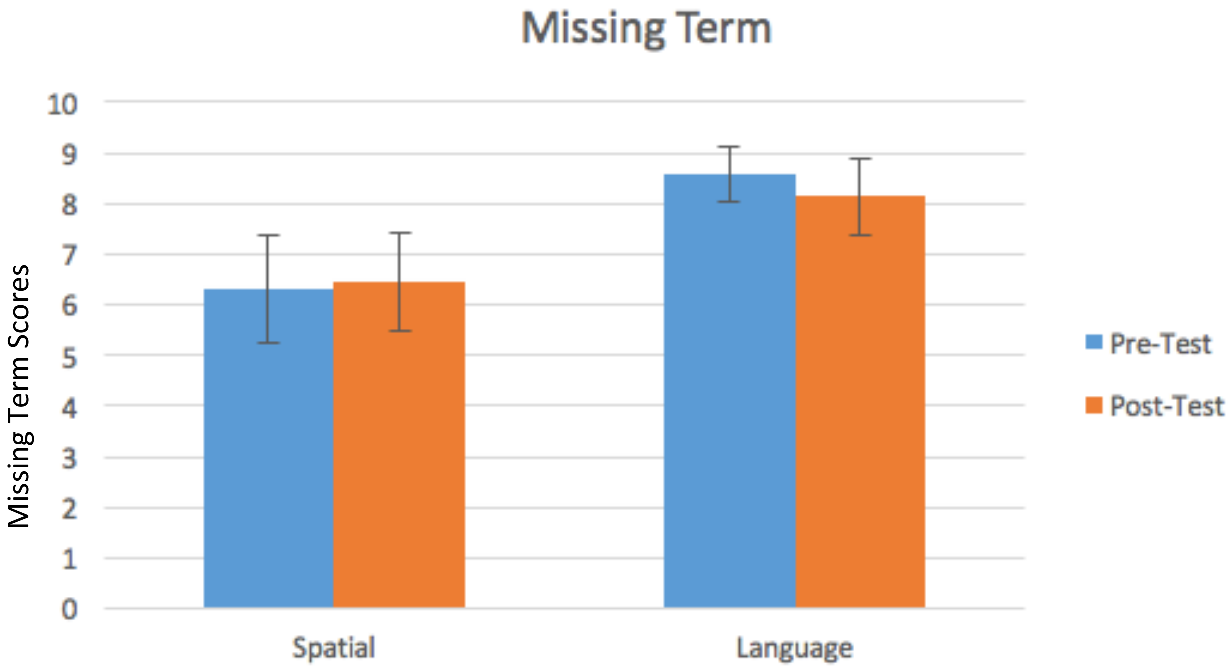

Figure 5. The pre-test and post-test of the missing term scores for both Spatial group and Language group. Error bars represent standard error of the mean.

Appendix E

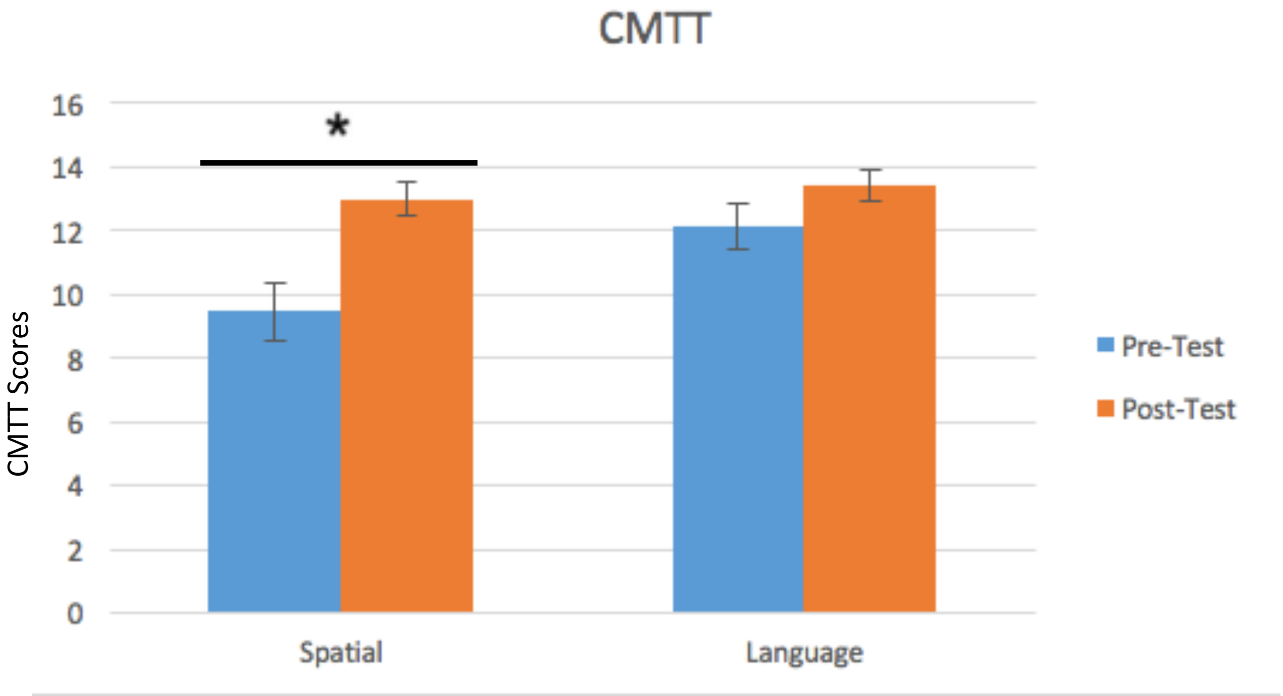

Figure 6. The pre-test and post-test of the CMTT scores for both Spatial group and Language group. Error bars represent standard error of the mean and the * represents statistically significant differences.

Appendix G

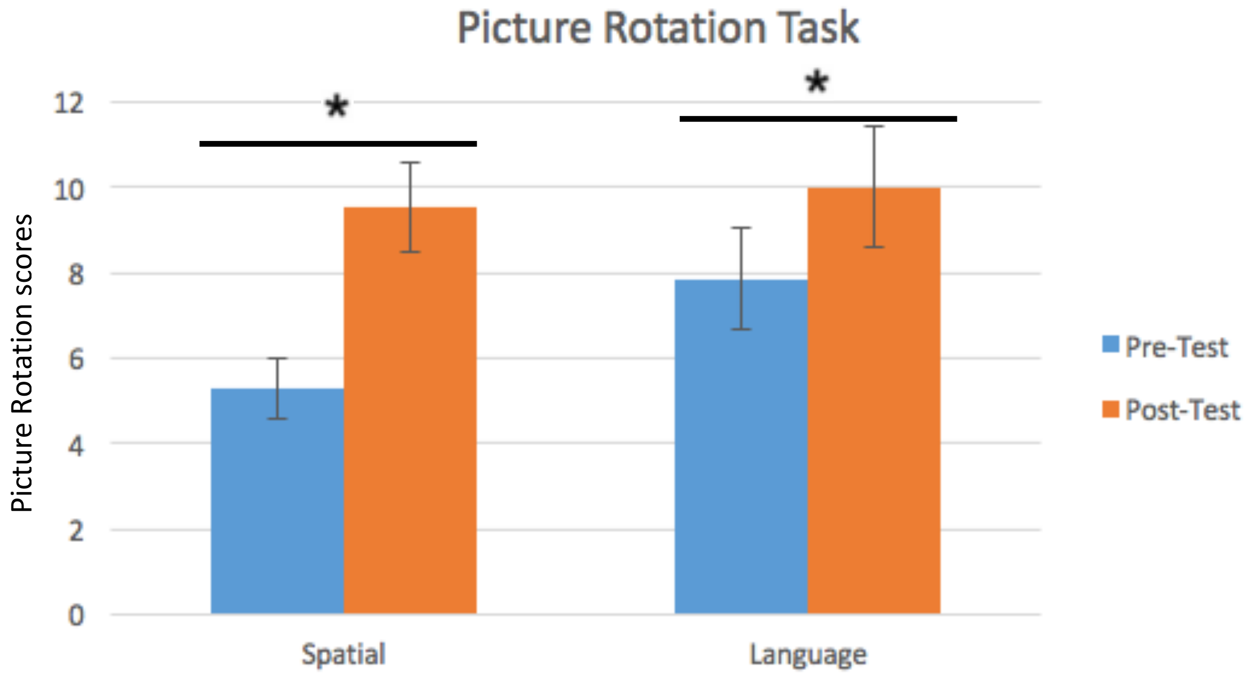

Figure 7. The pre-test and post-test of the Picture Rotation Task scores for both Spatial group and Language group. Error bars represent standard error of the mean and the * represents statistically significant differences.

Appendix H

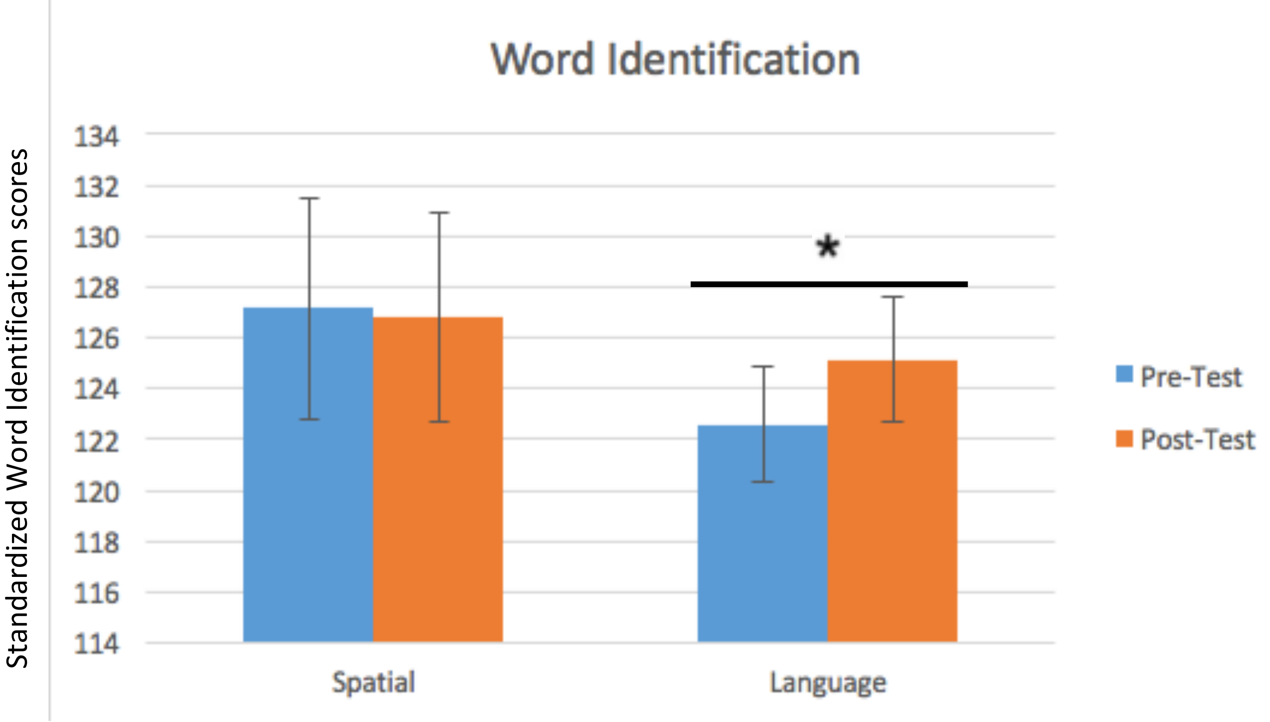

Figure 8. The pre-test and post-test of the Word Identification scores for both Spatial group and Language group. Error bars represent standard error of the mean and the * represents statistically significant differences.

Appendix I

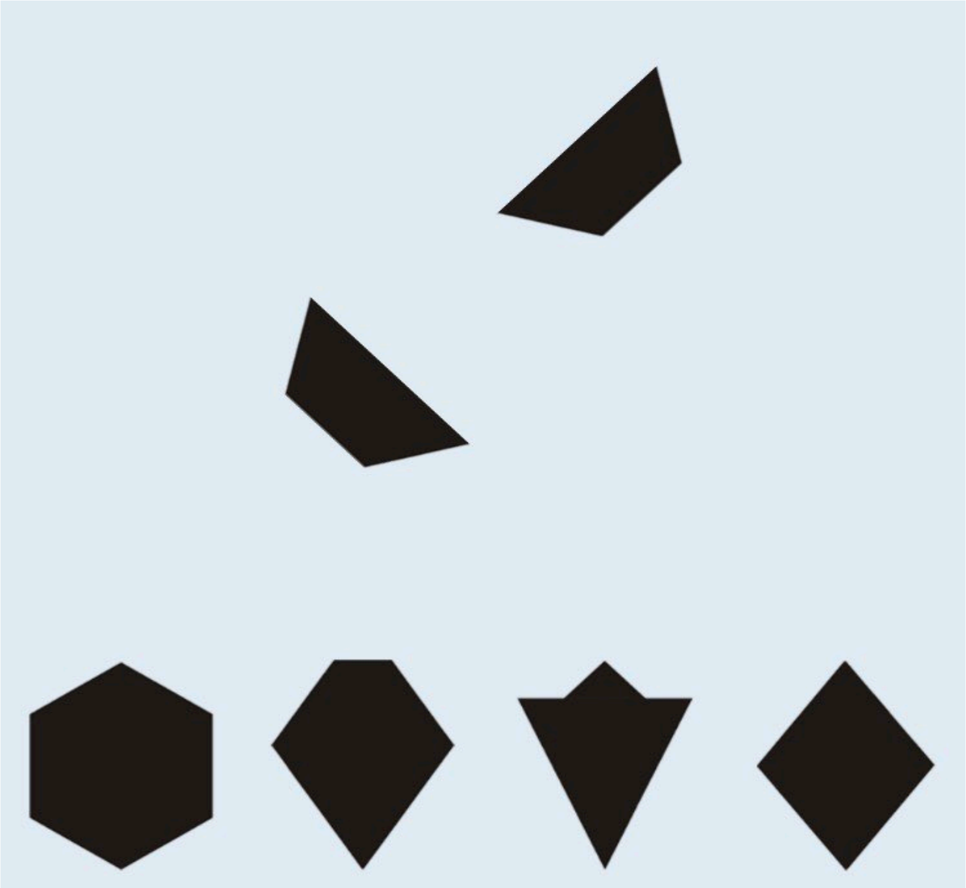

Figure 9. The layout of the spatial training website. The two pieces on top are the stimuli and the participants are to choose which shape the two pieces on top make.

Appendix J

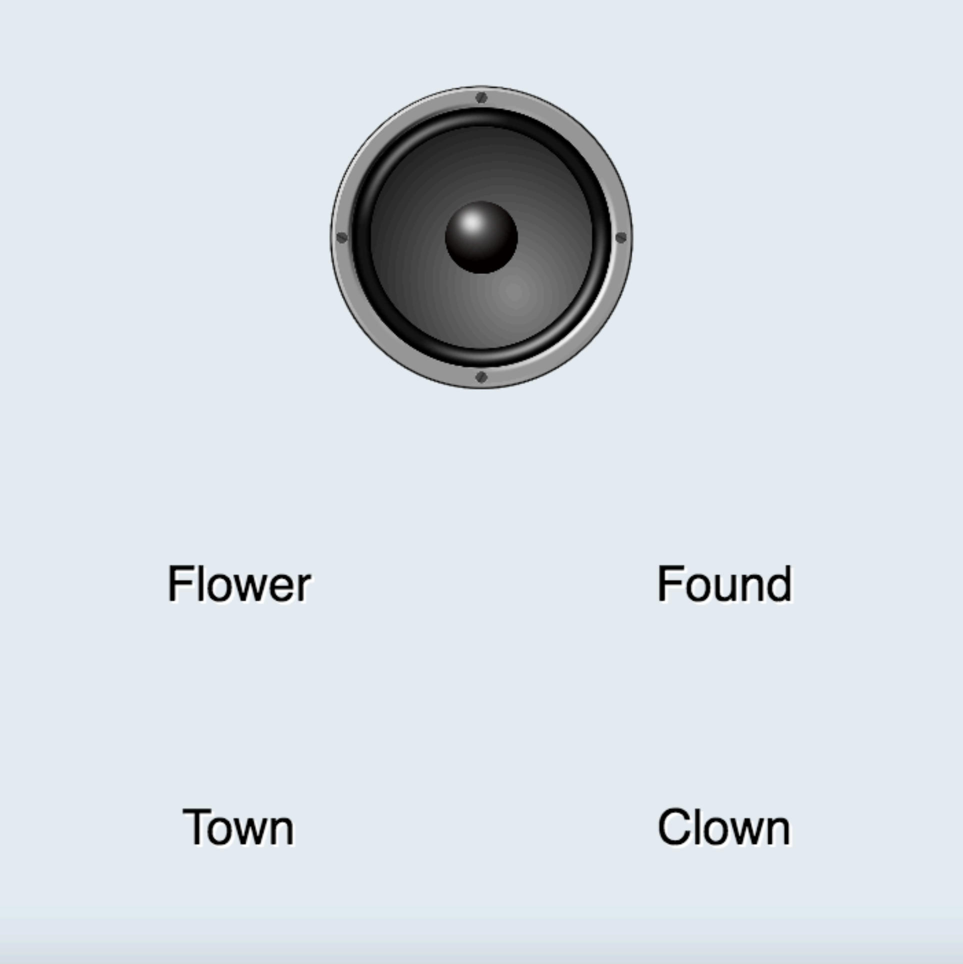

Figure 10. The layout of the language training. Participants can tap on the speaker to hear the target word as many times as they need. They are to choose the correct answer out of the four choices.

Appendix K

1 + ______ = 4	______ + 4 = 5
______ + 2 = 6	7 + ______ = 13
8 + ______ = 17	______ - 1 = 1
4 - ______ = 3	12 + ______ = 16
______ - 7 = 2	______ + 11 = 24
15 - ______ = 5	______ - 5 = 13

Figure 11. The missing term problems worksheet. The first question is a practice question and the remaining questions are scored 1 point each.