Cognition in Action: The relation between physical and mental paper folding in young children

David Menendez (damenend@ucsc.edu)
Department of Psychology, 1156 High Street Santa Cruz, CA 95060 USA
Samuel R. Halama (halama@wisc.edu)
Taylor M. Johnson (taylorjohnsonx3@gmail.com)
Department of Psychology, 1202 W. Johnson Street, Madison, WI 53706 USA
Karl S. Rosengren (kroseng2@UR.Rochester.edu)
Department of Psychology & Department of Brain & Cognitive Sciences, Meliora Hall
Rochester, NY 14627 USA

Abstract

Physically folding paper is a common activity performed by many children, but it is not mastered until middle-childhood. Paper folding ability has been the focus of studies motor development. There has been a long history in cognitive science of assessing spatial skills through mental paper folding tests. Despite the similarities between physical and mental paper folding, it is currently unknown whether there is a relation between physically and mentally folding paper. This study examined 107, 3- to 8-year-old children in both skills. Our results show that children of all ages were able to physically fold paper, but became more accurate with age. Additionally, we found that there is a significant relation between physical and mental paper folding, and that this relation was robust to different statistical controls and statistical specifications. We discuss how these results influence our understanding of the co-development of cognitive and motor skills.

Keywords: Motor development; Spatial skills; Paper folding

Introduction

Mental paper folding tasks have been a common way to assess spatial ability (Ekstrom, French, Harman, & Dermen, 1976; Uhlaner, Bolanovich, et al., 1952). Part of the interest for this and other visuo-spatial tasks comes from findings that spatial skills are associated with attainment and success in science, technology, engineering, and mathematics (STEM) disciplines (Humphreys, Lubinski, & Yao, 1993; Shea, Lubinski, & Benbow, 2001; Uttal, Miller, & Newcombe, 2013; Wai, Lubinski, Benbow, & Steiger, 2010). This work has highlighted the importance of spatial ability for STEM success. Test of mental paper folding are available for 4- 7-year-old children, allowing researchers to measure these skills in early childhood (Harris, Newcombe, & Hirsh-Pasek, 2013). In tandem, work on motor development has been examining similar questions with tasks that require physical (rather than mental) paper folding. Physical paper folding has been used as a measure of fine motor skills (Bruininks & Bruininks, 2005), and fine motor skills have also been linked to educational outcomes such as performance on science, mathematics and readings tasks in pre-school children (Grissmer, Grimm, Aiyer, Murrah, & Steele, 2010). The similarity between physical paper folding (i.e., manipulating paper to create intended folds) and mental paper folding (i.e., manipulating a mental representation of paper to predict how it would look after it is folded) is hard to deny, as they are analogous to one another and the work described so far shows that they predict similar outcomes. However, because of the lack of conversation between the cognitive and motor development literature it is currently unknown whether physical and mental paper folding abilities are related to one another. In this exploratory study, we investigated this relation among 3- to 8-year-old children.

Mental paper folding

Mental paper folding tasks assess a person's ability to perform dynamic spatial transformations to mental representation. Other tasks that assess dynamic spatial transformations include mental rotation tasks (Thurstone, 1951). Traditional work on mental rotation with children suggest that children perform at chance until the age of 7 (Dean & Harvey, 1979), but more recent work suggests that children as young as 4 perform above chance in a simplified mental rotation task (Levine, Huttenlocher, Taylor, & Langrock, 1999). Work on children's mental paper folding abilities is more limited, but suggest that children can perform above chance by 5.5 years of age (Harris et al., 2013), and that this ability continues to improve with age (Ang & Lee, 2010).

Physical paper folding

Physical paper folding is a common practice in many communities, often included in making crafts, wrapping presents and more dedicated practices like origami. Classroom observations in the United States indeed show that this is a common activity for elementary school children (McHale & Cermak, 1992). Paper folding has also been used in developmental psychology studies as a way to investigate attention and learning in the United States, but also indigenous communities, such as Guatemalan Mayan communities (Correa-Chávez, Mejía-Arauz, & Rogoff, 2015; Correa-Chávez & Rogoff, 2009). In a study on the development of physical paper folding skills Travers et al. (2018) had children (ages 18 months to 7 years) fold paper of different shapes. They found that children as young as 27 months of age could produce measurable folds, and their skills increased with age. It wasn't until ages 4 and 5 that over half of the children were completing high quality folds. The researchers also measured the distance between the target fold that children were tasked with and their actual fold, and found that with age, children became more accurate. They also found that children made mistakes such as always folding the paper in half, folding the opposite corner, or folding the paper to the folding line (rather

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

than folding on the line). Therefore, work on physical paper folding shows that this task can be used among children in different communities, with very young children (younger than most mental paper folding tasks), and that it captures a high degree of individual differences.

Relation between physical and mental paper folding

At first glance it might seem obvious that mental and physical paper folding performance would be related. However, it is important to note that these tasks are not identical. In physical paper folding, children are able to plan their movements and receive immediate visual feedback on their performance. Additionally, as described previously, physical paper folding is common for many children, while mental paper folding is not. Mental paper folding is a more abstract and unfamiliar task for children. Also, the fine motor control required to physically fold paper can be a challenge for many children as these skills are still developing (Travers et al., 2018). Therefore, their relation might not be as trivial as it appears.

Training studies have provided mixed evidence about how paper folding training affects mental paper folding skills. Embertson (1987) had adults fold paper for 15 minutes or complete a filler task for the same amount of time. Afterwards, all participants completed a mental paper folding task. They found that participants who folded paper had better mental paper folding scores than those who did the filler task. Some of these studies suggest that origami practice enhances spatial skill in 4th through 7th graders (Cakmak, Isiksal, & Koc, 2014; Krisztián, Bernáth, Gombos, & Vereczkei, 2015; Taylor & Hutton, 2013), but other work has not found these effects (Boakes, 2009). Only one of these studies (Cakmak et al., 2014) had participants complete both mental and physical paper folding tasks at baseline, but they do not report whether these two tasks were related. Therefore, limiting the generalization we can make from these data.

Current study

In the current study, we were interested in the relation between mental and physical paper folding. We examined this among 3- to 8-year-old children. We hypothesized that there will be a relation between these tasks, such that children who are better at physically folding paper will also be better at mental paper folding.

Method

Participants

We recruited 107 3- to 8-year-old children (49 boys, 56 girls, and the parents of 2 children did not report gender) to complete the study. Children completed the study either at a local children's museum (n = 94) or at the laboratory in a large mid-Western university (n = 13). Data from three children was excluded because they did not finish the study. Children who completed the study at the children's museum received a small toy for their participation. Children who completed the study in the lab receive $15 dollars. Parental consent and

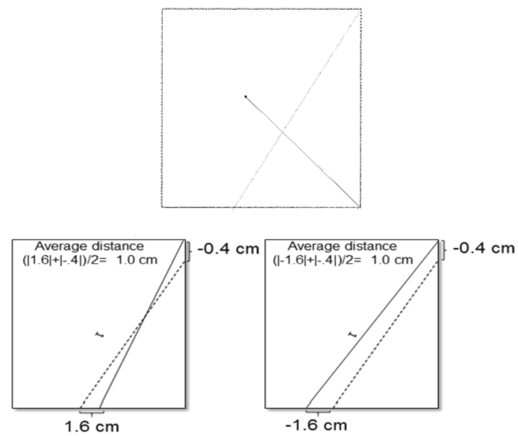

Figure 1: Top image shows a sample paper folding trial. The dotted line shows the target fold, and the arrow the direction in which the paper needs to be folded. Bottom images show how we measured the accuracy of the folds, with the solid line showing the actual fold a child made.

child assent were collected before the start of the study. The study was approved by the local IRB.

Materials

Physical paper folding Our physical paper folding task was modeled after the task by Travers et al. (2018). We had children perform four different folds on three shapes (rectangles, triangles, and diamonds). The paper was cut to the appropriate shape, and the intended fold was printed on the paper. The order of the shapes and the order of the folds within a shape was randomized for each participant. See Figure 1 for an example. Unlike the mental paper folding task (see details below), the physical paper folding was done on white paper (as having two colored sheets of paper together would have made the paper thicker and increased the difficulty of folding the paper). Based on our coding of folding quality (see coding section), this measure had good internal consistency (Cronbach's $\alpha = 0.96$).

Mental paper folding We created our own mental paper folding task modelled after Harris et al. (2013). The children were shown a paper that was blue in one side and red on the other. Then, all children were shown the same physical fold of a rectangle to demonstrate how folding the blue side would reveal the red side. Then, they started the mental paper folding task. In this task, children were shown a fold and were given either two or three response options (with the correct fold always being one of the options). The incorrect answer options could be either an over fold (i.e., folded too much), an under fold (i.e., folded too little), or a wrong fold (i.e., the incorrect side was folded). When the question had

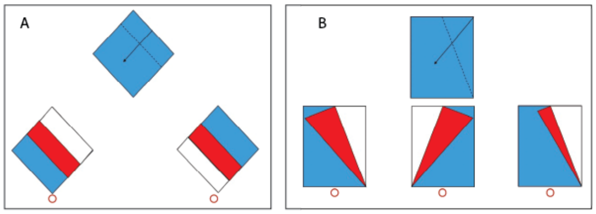

Figure 2: Sample items from mental paper folding task. Panel A shows an example of the 2-choice questions for the diamond. Panel B shows an example of the 3-choice questions for the rectangle

two answer options, the incorrect answer would be one of these three errors. When the question had three answer options, there were two different types of errors. For each type of question, children were asked about same three shapes as in the physical paper folding task: rectangles, triangles, and diamonds. They were shown four different folds per shape per question, leading to a total of 24 folds. The folds were not the same for the two and three answer options questions. The two answer option questions were always presented first. The order in which the shapes were presented, the order of the folds within a shape, and the order of the answer options were randomized for each child. This measure had good internal consistency (Cronbach's $\alpha = 0.86$).

Procedure

Children completed the mental paper folding and the physical paper folding tasks in one session. Which task was completed first was counter-balanced across participants. The entire study took less than 20 minutes. To help with motivation during the mental paper folding task, children were told that they were helping the local university's mascot solve puzzles. After children answered all the questions for one shape, they would see a picture of the mascot celebrating and telling them that they were doing great (independent of their actual performance).

Coding

For the mental paper folding task, we counted number of correct answers. For the physical folds, we used the distance (in millimeters) between the child's fold and the expected fold as a measure of accuracy. As can be seen in Figure 1, we took the average of the deviation between the two end points of the child's fold.

Additionally, we coded children's physical folds into six categories: correct folds, folds to line, folds in half, wrong corner folds, and crumbled. Folds to line was when the child folded the paper up until the line, so that the edge of the paper touched the printed line showing the expected fold. Folds in half were when the child folded the paper in half. We divided these folds between the ones in which the child's fold was perpendicular to the expected fold, and the ones in which it was not. We made this distinction because we could not calculate a distance score for the perpendicular folds. Wrong corner folds were folds in which the child folded the incorrect corner or side. Crumbled were trials in which the child rolled or crumbled the paper, not producing any folds. One coder categorized all of the folds. To check for reliability, a second coder categorized all of the folds made by 25 children (23.4% of the total sample), and disagreements were resolved through discussion. The reliability between their codes was considered acceptable (Cohen's $\kappa = 0.74$).

Results

We first present the results for physical paper folding, then metal paper folding. Then we examine the relation between the two.

Physical paper folding

To examine children's accuracy in completing physical folds, we calculated the average deviation between the child's fold and the expected fold. On average, children deviated 15.22 mm (SD = 16.99mm) from the expected line. We fit a linear mixed-effects model predicting average distance from children's age (mean-centered), gender, task order, and the shape being folded (as fixed effects). We included by-subject random intercepts and by-subject random slopes for the effect of shape. For the effect of shape, we used rectangles as the reference group. We found that children's folding accuracy differed by shape, $F(2, 93.89) = 5.56$, $p = .005$. When children folded rectangles (M = 16.90, SD = 18.85) they had higher deviations (i.e., were less accurate) than when they folded diamonds (M = 14.54, SD = 17.23) or triangles (M = 14.22, SD = 14.51), $t(456.93) = 2.85$, $p = .005$ and $t(90.05) = 2.90$, $p = .005$ respectively. Deviations for triangles and diamonds did not differ from one another, $p = .533$.

Additionally, we found that age was related to average deviation, with older children having smaller deviations, $F(1, 97.90) = 42.63$, $p < .001$. We also found that participants who completed the physical paper folding task first (M = 16.55, SD = 17.26) had higher deviation scores than those who completed the mental paper folding task first (M = 13.58, SD = 16.51), $F(1, 95.24) = 4.56$, $p = .035$. We did not find any gender differences in physical paper folding, $F(1, 94.31) = 0.01$, $p = .949$.

We did see that there were differences in the quality of children's folds. The 107 children produced 1273 folding attempts. Of these attempts, 2.20% (28 folds) of the attempts the children just crumbled or rolled the paper, not producing any folds. In 11.78% (150 folds) of the attempts the children folded the paper in half. Of these fold in half trials, 61 (40.67% of the folds in half) were perpendicular to the expected fold so we could not calculate a distance. For the remaining folds in half (59.33%) we could calculate a deviation from the expected line. Of the original folds, in 1.81% (23 folds) of the attempts the children folded the wrong side or corner. In 7.77% (99 folds) of the attempts the children folded the paper up to the line (instead of over the line). The

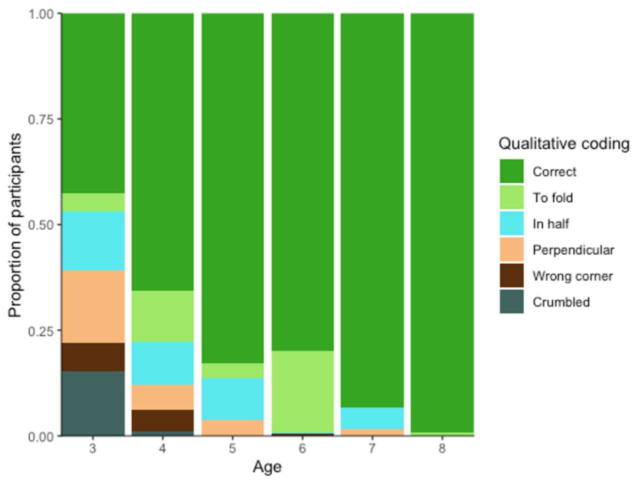

Figure 3: Distribution of folds by age and quality codes.

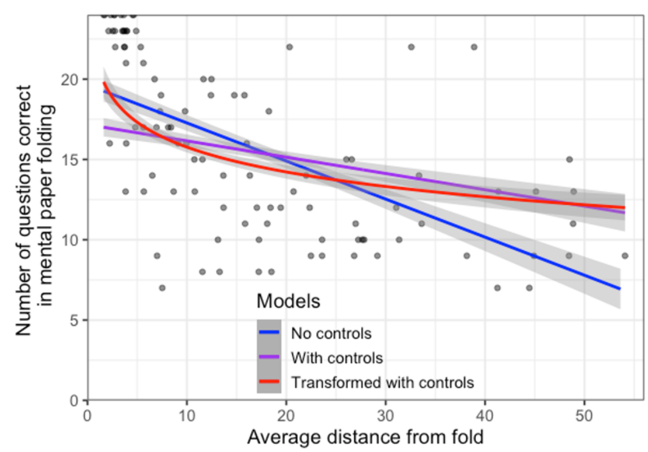

Figure 4: Relation between physical paper folding (in the x-axis) and mental paper folding (in the y-axis). We plot three lines: one for the model with no covariates, one controlling for age, gender and order, and one where distance was log (base 2) transformed controlling for age, gender and order. Error bands show the between subject standard error of the point estimate.

remaining folds (76.4% or 973 folds) were coded as correct folds.

We used a chi square test to examine whether the quality of folds depended on children's age. We did find a significant relation, $\chi^2(25, N = 1{,}237) = 447.09$, $p < .001$. As can be seen in Figure 3, the folds of older children were more likely to be coded as correct. This trend is seen across all shapes.

Mental paper folding

The mental paper folding tasks was divided into questions with two response choices followed by questions with three answer choices. On average children got 70% of the two answer questions correct and 61% of the three answer questions correct. The reliability of the mental paper folding task for both types of questions together was acceptable (α = .86), suggesting that we could combine all the questions into one measure.

We fit a linear mixed-effect model to predict the probability that children answered a question of the mental paper folding task correctly. We included shape (with rectangles as the reference group), age (mean-centered), gender and order as fixed effects. We also included by-subject random intercepts and by-subject random slopes for the effect of shape. We only found an effect of age, such that older children made more correct choices than younger children, $F(1, 93.55) = 126.89$, $p < .001$. No other effects were significant.

Relation between physical and mental paper folding

To examine the relation between physical and mental paper folding, we used children's average deviation scores for the physical paper folding and the number of correct responses for mental paper folding. We fit a linear regression predicting number of mental paper folding questions answered correctly from average deviation from expected physical fold. As hypothesized, we found that the smaller the average deviation (i.e., the more accurate the physical fold) the more metal paper folding questions children answered correctly, b = -0.24, $F(1, 102) = 60.19$, $p < .001$. We then included age, gender, and order as covariates, to examine if these relation would persist. The relation between physical and mental paper folding was still significant, b = -0.10, $F(1, 94) = 11.45$, $p = .001$. Age was also a significant predictor of mental paper folding, $F(1, 94) = 69.07$, $p < .001$. No other effects were significant. The effects were the same if we fit the a linear mixed-effects model predicting the probability of answering each mental paper folding question correctly, $F(1, 96.32) = 54.12$, $p < .001$ with no controls, $F(1, 93.84) = 11.40$, $p < .001$ after controlling for age, gender and order. Figure 5 shows the relation between physical and mental paper folding both with and without covariates.

We also checked if we would still see the same relation if we examined only correct folds. When controlling for age, gender, and order, we still saw an effect of average of physical folds on the number of mental paper folding questions answered correctly, b = -.21, $F(1, 91) = 24.64$, $p < .001$. The relation did not change if we used a linear mixed-effect model, $F(1, 98.49) = 20.98$, $p < .001$. We also checked for violation of model assumptions for the model with all the folds. We found that there was a violation of the linearity assumption. We used a log base 2 transformation on average deviation from the fold. After this transformation, all the assumptions of the general linear model were met. We still found the same relation between mental and physical paper folding, $F(1, 94) = 23.00$, $p < .001$.

We then compared the proportion of mental paper folding questions that children answered correctly to the number of folds they made that were coded as correct. We fit a mixed-effects linear regression predicting the proportion correct from type of test, age and their interaction (with by-subject random intercepts). Again, we found that older children did better on both tests than younger children, $F(1, 95.54) = 81.45$, $p < .001$. Critically, we found that children had higher scores in the physical ($M = 0.78$, $SD = 0.33$) than the mental ($M = 0.65$, $SD = 0.23$) paper folding tasks, $F(1, 100.29) = 18.44$, $p < .001$. Their interaction was not significant.

Discussion

In this study, we were interested in the relation between two common tasks, mental and physical paper folding. Mental paper folding has been used as a way to measure spatial skills, particularly dynamic spatial transformation abilities. Physical paper folding has been used to assess fine motor skills. Given the disconnect between the motor and cognitive development literature, there has been no work examining the relation between these tasks. We examined 3- to 8-year-old children in both of these tasks, and found the hypothesized relation. Children who made more accurate folds also performed better on a mental paper folding task. This relation persisted across various model specifications, including controlling for age and gender (two known predictors of spatial skills), thus suggesting that this relation is not merely because older children are better at both tasks.

Given the exploratory nature of this study, it is not possible for us to ascertain the nature of this relation. It is possible that this relation exists because children use their experience folding paper to guide their decisions in mental paper folding tasks. This would suggest that physical paper folding may serve as a building block of mental paper folding, a notion similar to Piaget (1954) idea that cognitive capabilities stem from early sensorimotor experiences. This unidirectional causality would suggest that experience folding paper (and fine motor control) will lead to better spatial skills. Given the correlational nature of our data, we cannot establish this unidirectionality. Another possibility is that the same spatial skills that children utilize during the mental paper folding task, are also being tapped during physical paper folding. This would suggest that the relation is due to both tasks recruiting the same cognitive mechanisms. This is akin to latent construct models that are popular in other areas of cognition, such as IQ (Van Der Maas et al., 2006). So, as spatial skills improve, performance on both physical and mental paper folding would similarly improve. It is unlikely that this model is entirely accurate, as physical paper folding also requires motor skills that may not improve at the same rate as spatial skills. However, this model would be entirely consistent with our findings.

Another possible model that we would like to advance relies on the relation between cognitive and motor development (Needham, Barrett, & Peterman, 2002). The co-development of cogntive and motor systems could produce positive relations across both systems, and furthering development across both. This reciprocal causation (also called mutualism) has been seen across other areas of cognition (Van Der Maas et al., 2006), and development (Rittle-Johnson, 2017), and is in accordance with Dynamic Systems theories of development (Perone, Simmering, & Buss, 2021; Smith & Thelen, 2003). Physical paper folding requires both spatial abilities and fine motor skills. The feedback that a child receives upon folding the paper leads them to refine both skills. When children engage in other activities that require spatial skills, these skills improve, which then might lead them to be more accurate in future physical paper folds. While we cannot distinguish between the latent construct and reciprocal causation models, both of these models align with common coding theory that suggests that similar representations are activated when a motor plan is perceived or actively performed (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1997). Whether this common code is present early on in development or emerges from experiences needs to be explored in future work that examines the longitudinal relation between motor and cognitive development.

This study also advances our understanding of physical paper folding. Our results replicate prior findings showing that children can complete folds in early childhood, but they do not consistently do so until age 4 (Travers et al., 2018), and we extend this to 8-year-old children. We also characterized the errors that children made in their folding, and showed how the amount and type of errors decrease with age. Folding the wrong side of the paper is rare after age 4. However, folding the paper in half or folding to the fold line persisited until age 7, with minimal errors in age 8. Additionally, the relation between mental and physical paper folding suggests that future work could use physical paper folding task to measure spatial skills, which can be administered to children as young as 27 months of age (Travers et al., 2018). We found that scores on the physical paper folding task were higher than the mental paper folding task. But this might have been because the correct fold codes for the physical task did not consider over or under folds. In the mental task, some of the incorrect answers included images where the fold was over extended (folded too much paper) or under extended (folded too little). But this was not taken into account when coding the physical folds as correct or not (given that the distance from the fold was also measured). Therefore, the difference between these two tasks should be interpreted with caution.

This study should be understood in light of its limitations. First, this study was conducted in the United States where physical paper folding is common. Although paper folding tasks have been used in different cultures (Correa-Chávez & Rogoff, 2009; Correa-Chávez et al., 2015), how children interpret pictures like the ones used in the mental paper folding test varies by culture (Zhu et al., 2025). Therefore, it is possible that the results will vary depending on children's familiarity with both tasks. Second, the two tasks were not completely analogous. While the same shapes were used in both tasks, the mental paper folding showed a paper that was blue on one side and red on the other, while the physical paper folding task use white sheets of paper. Making a paper that was blue on one side and red on the other would have meant having a thicker piece of paper that was more difficult for children to fold. However, this decision did decrease the similarity of the task. It is possible that having the different colors might have helped children understand the mental paper folding task, or could have also introduced unnecessary complexity to the visuals. Future research should investigate these possibilities. Additionally, we were not able to ask families several demographic questions due to restrictions with collecting data at the children's museum. As such, we cannot say whether other demographic factors (such as SES), influence the results.

Overall, this study shows that there is a relation between physical and mental paper folding. There are several possible explanations for this relation, but our results suggest that there could be a dynamic interplay between cognitive and motor development that promotes the development of both skills.

References

Ang, S. Y., & Lee, K. (2010). Exploring developmental differences in visual short-term memory and working memory. *Developmental Psychology*, *46*(1), 279-285.

Boakes, N. J. (2009). Origami instruction in the middle school mathematics classroom: Its impact on spatial visualization and geometry knowledge of students. *RMLE Online*, *32*(7), 1–12.

Bruininks, R. H., & Bruininks, B. B. (2005). *Bruininks-oseretsky test of motor proficiency, second edition (bot2)*. Minneapolis, MN: Pearson Assessment.

Cakmak, S., Isiksal, M., & Koc, Y. (2014). Investigating effect of origami-based instruction on elementary students' spatial skills and perceptions. *The Journal of Educational Research*, *107*(1), 59–68.

Correa-Chávez, M., Mejía-Arauz, R., & Rogoff, B. (2015). Children learn by observing and contributing to family and community endeavors: A cultural paradigm. *Child Development*, *49*(3), 664-678.

Correa-Chávez, M., & Rogoff, B. (2009). Children's attention to interactions directed to others: Guatemalan mayan and european american patterns. *Developmental Psychology*, *45*(3), 630-641.

Dean, A. L., & Harvey, W. O. (1979). An information-processing analysis of a piagetian imagery task. *Developmental Psychology*, *15*(4), 474-475.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (Eds.). (1976). *Manual for kit of factor-referenced tests*. Princeton, NJ: Educational Testing Service.

Embertson, S. E. (1987). Improving the measurement of spatial aptitude by dynamic testing. *Intelligence*, *11*(4), 333–358.

Grissmer, D., Grimm, K. J., Aiyer, S. M., Murrah, W. M., & Steele, J. S. (2010). Fine motor skills and early comprehension of the world: two new school readiness indicators. *Developmental Psychology*, *46*(5), 1008-1017.

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

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (tec): A framework for perception and action planning. *Behavioral and brain sciences*, *24*(5), 849–878.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology*, *78*(2), 250-261.

Krisztián, Á., Bernáth, L., Gombos, H., & Vereczkei, L. (2015). Developing numerical ability in children with mathematical difficulties using origami. *Perceptual and Motor Skills*, *121*(1), 233–243.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*(4), 940-949.

McHale, K., & Cermak, S. A. (1992). Fine motor activities in elementary school: Preliminary findings and provisional implications for children with fine motor problems. *The American Journal of Occupational Therapy*, *46*(10), 898–903.

Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky mittens' enhances young infants' object exploration skills. *Infant Behavior and Development*, *25*(3), 279–295.

Perone, S., Simmering, V. R., & Buss, A. T. (2021). A dynamical reconceptualization of executive-function development. *Perspectives on Psychological Science*, *16*(6), 1198–1208.

Piaget, J. (1954). *The construction of reality in the child*. Routledge.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154.

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives*, *11*(3), 184–190.

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.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343–348.

Taylor, H. A., & Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to stem education. *Cognition and Instruction*, *31*(4), 434–455.

Thurstone, L. L. (1951). Some primary abilities in visual thinking. *Proceedings of the American Philosophical Society*, *94*(6), 517–521.

Travers, B. G., Kirkorian, H. L., Jiang, M. J., Choi, K., Rosengren, K. S., Pavalko, P., & Jobin, P. (2018). Knowing how to fold 'em: Paper folding across early childhood. *Journal of Motor Learning and Development*, *6*(1), 147–166.

Uhlaner, J., Bolanovich, D., et al. (1952). Development of the armed forces qualification test and predecessor army screening tests, 1946–1950. *PRB report*.

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.

Van Der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842-861.

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (stem) and its relation to stem educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, *102*(4), 860-871.

Zhu, R., Pitchik, H. O., Kilonzo, T. N., Engelmann, J., Fernald, L. C., & Gopnik, A. (2025). The development of picture comprehension across early environments: Evidence from urban and rural toddlers in western kenya. *Developmental Science*, *28*(1), e13579.