

Shape Up: An Eye-Tracking Study of Preschoolers' Shape Name Processing and Spatial Development

Brian N. Verdine
University of Delaware

Ann Bunger
Indiana University

Angeliki Athanasopoulou
and Roberta Michnick Golinkoff
University of Delaware

Kathy Hirsh-Pasek
Temple University

Learning the names of geometric shapes is at the intersection of early spatial, mathematical, and language skills, all important for school-readiness and predictors of later abilities in science, technology, engineering, and mathematics (STEM). We investigated whether socioeconomic status (SES) influenced children's processing of shape names and whether differences in processing were predictive of later spatial skills. Three-year-olds ($N = 79$) with mothers of varying education levels participated in an eye-tracking task that required them to look at named shapes. Lower SES children took longer to fixate target shapes and spent less time looking at them than higher SES children. Gaze variables measured at age 3 were predictive of spatial skills measured at age 5 even though the spatial measures did not require shape-related vocabulary. Early efficiency in the processing of shape names may contribute to the development of a foundation for spatial learning in the preschool years.

Keywords: language, geometry, spatial, eye-tracker, preschool

Supplemental materials: <http://dx.doi.org/10.1037/dev0000384.supp>

Early geometry and spatial skills predict success in school, especially in the so-called STEM subjects (science, technology, engineering, and mathematics; Mix & Cheng, 2012; Wai, Lubinski, Benbow, & Steiger, 2010). The term *spatial skills* refers to the encoding and mental manipulation of information about the configuration of objects and their location in our environment (Uttal et al., 2013). With recent changes in early curricula, learning about

shapes and spatial skills both play a major role in school readiness for mathematics (e.g., [Common Core State Standards Initiative, 2010](#); [Office of Head Start, 2011](#)). Understanding early STEM-related learning is, therefore, critical for improving school readiness and increasing later STEM achievement, and may be essential for minimizing socioeconomic status (SES) achievement gaps in STEM-related domains (e.g., [Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2017](#)).

This study investigated 3-year-old children's processing of shape names in a mixed-SES sample. The same children were then given tests of spatial skills over two more years to investigate connections between shape name processing and later spatial skills. The study had two main goals: (a) to assess whether differences in maternal education influence the time course of the identification of named shapes, and (b) to determine whether individual differences in shape name processing predict future spatial skills.

Early Language Experience

Differences in early experiences can have important implications for both vocabulary acquisition and language processing speed. For example, children from lower SES households appear to have difficulty quickly learning new words ([Horton-Ikard & Weismer, 2007](#)). One explanation is that children in lower SES households receive less and lower quality language input than those from higher SES families (e.g., [Hart & Risley, 2003](#); [Hirsh-Pasek et al., 2015](#)). Additionally, children from lower SES families often attend lower quality childcare programs where language

This article was published Online First July 31, 2017.

Brian N. Verdine, School of Education, University of Delaware; Ann Bunger, Department of Linguistics, Indiana University; Angeliki Athanasopoulou, Department of Linguistics and Cognitive Science, University of Delaware; Roberta Michnick Golinkoff, School of Education, University of Delaware; Kathy Hirsh-Pasek, Department of Psychology, Temple University.

This research was supported by the National Institutes of Health Stimulus Grant 1RC1HD0634970-01, the National Science Foundation via the Spatial Intelligence and Learning Center (Grant SBE-1041707), and the Institute of Education Sciences through Grant R305A140385.

We thank Nora Newcombe, Nancy Jordan, and Marcia Halperin for their consultation on this project. Wilkey Wong and Andrew Filipowicz were also essential in getting this project off the ground. We would also like to thank the Spatial Intelligence and Learning Center and their Spatial Network for their assistance at various stages of this project.

Correspondence concerning this article should be addressed to Brian N. Verdine or Roberta Michnick Golinkoff, School of Education, University of Delaware, 224 Willard Hall, Newark, DE 19716. E-mail: brian.verdine@gmail.com or roberta@udel.edu

development is not emphasized compared with those attended by children from higher SES families (Magnuson, Meyers, Ruhm, & Waldfogel, 2004). These differences in early experience may contribute to the SES gap in language development.

The amount and complexity of parental language input at 18 months can predict not only children's vocabulary size at 24 months, but also the speed with which they resolve word meaning (Hurtado, Marchman, & Fernald, 2008). Differences in processing speed are important because they are likely to influence comprehension broadly. Understanding the speech stream requires listeners to continuously process language at multiple levels (e.g., from phonemes to phrases to clauses), and to integrate acoustic information with linguistic and conceptual knowledge to interpret language as it unfolds (Fernald & Weisleder, 2011). Children who are slower in understanding speech may need to exert more effort in trying to keep up with the speech stream, effectively limiting the cognitive resources left to comprehend less-familiar words. The effects of slower language processing may not be evident in everyday contexts in which children are processing familiar words and structures in rich environmental contexts. However, they may be magnified in contexts like STEM learning, in which children are likely to encounter unfamiliar words that would be expected to take more time to process and more cognitive resources to comprehend.

Preschool Shape Learning

Shape categories are differentiated by their spatial or mathematical properties, for example, the number of sides, sizes of sides or angles, and symmetry. To quickly identify a shape after hearing its name, children must not only be able to recall some mental representation that shares the shape's defining spatial and numerical properties, but also to process the shape name quickly and map it onto the mental representation.

Young children begin to show interest in shapes while they are still quite young (Shutts, Örnkloo, von Hofsten, Keen, & Spelke, 2009) and begin to accurately label them during the preschool years, with some variation across shapes and understanding of shapes' defining properties (e.g., Satlow & Newcombe, 1998). Between 25 and 30 months, children begin identifying standard versions of common shapes (Verdine, Lucca, Golinkoff, Hirsh-Pasek, & Newcombe, 2016). By 3 years of age, over 60% of children can name basic shapes like circle, square, and triangle (Fuson & Murray, 1978) and those numbers climb near 80% by age 5 (Klein, Starkey, & Wakeley, 1999). Indeed, many parents of preschoolers would claim that their child "knows" their shapes. However, studies on early shape labeling have primarily used the most often encountered, equilateral versions and none that we are aware of have investigated *speed* in identifying shapes.

Opportunities to learn the range of exemplars in each shape category appear to also be relatively limited. Most children's toys focus on a small set of shapes (Resnick, Verdine, Golinkoff, & Hirsh-Pasek, 2016) that are mostly canonical. That is, they include prototypical versions which are especially salient even for adults (e.g., Rosch, 1973) and are usually equilateral or have some preferred ratio of side lengths (e.g., a "golden rectangle" with side ratios of approximately 1:1.62; Benjafield, 1976). In schools, educators infrequently use shape names (Rudd, Lambert, Satterwhite, & Zaier, 2008; Sarama & DiBiase, 2004), and rarely de-

scribe their properties when they do (Sarama & Clements, 2004). Research suggests that exposure to varied exemplars of a category is necessary to separate defining from incidental features (e.g., Rakison & Oakes, 2003). Given the relatively infrequent exposure to varied exemplars and the lack of description of their properties, it is perhaps not surprising that children need years of exposure to understand the features that define shapes (Satlow & Newcombe, 1998). At 30 months, children begin to extend shape labels to less standard versions, but they are not adept at doing so (Verdine et al., 2016).

The ability to identify shapes has been directly related to experience with shapes before formal instruction (e.g., Clements, Swaminathan, Hannibal, & Sarama, 1999). In addition, there is reason to believe that children's shape identification will be related to their language skills. It is well known that shape plays a role in early word learning. For example, children show a "shape bias" in learning object labels (Landau, Smith, & Jones, 1988) and, as they learn more object categories, they use shape to recognize stylized three-dimensional caricatures of known and novel objects (Smith, 2003). Given variability in how much and what kind of exposure children receive to shapes early in life, there may be individual differences in how efficiently they can process shape terms. These processing differences may be indicative of prior experience with shapes and instruction on their defining spatial properties, and may affect children's ability to learn about the spatial properties of shapes. For example, for children with good language skills and much experience with shapes and their names, some shapes may be identified quickly using lower-level, perceptual processes, regardless of changes in color or context. On the contrary, children with lower language skills and less experience with shapes and their names may experience processing bottlenecks that make it more difficult to identify shapes. In turn, these bottlenecks might make it more difficult to attend to new information about shapes, including considering their spatial properties. If this cascade of effects is accurate, the time course of children's identification of a set of prompted shapes should predict later spatial skills, even after accounting for the size of a child's vocabulary. Prior work does show that early experiences and exposure to language can influence spatial skills (Pruden, Levine, & Huttenlocher, 2011). Do children who are better at processing shape names also have better spatial skills?

The Current Study

Our Shape Fixation Task (SFT) incorporated an eye-tracker to observe the locus of children's attention as they processed shape labels for six categories of common geometric shapes (circle, triangle, square, rectangle, pentagon, and hexagon; Figure 1). With the SFT we explicitly measure children's ability to look to shapes in response to verbal prompts. Both the percent of time spent looking at targets and the speed with which they are selected are potentially meaningful indicators of how children process geometric language. Across trials, looking consistently more at target shapes than nontarget shapes requires some understanding of shape category labels and of the shared properties of shapes within a category. Prior work (Verdine et al., 2016) with 30-month-olds using an untimed comprehension test indicates that most children of this age would be at or near ceiling performance on the presented shapes. Our priority was to go beyond a simple accuracy

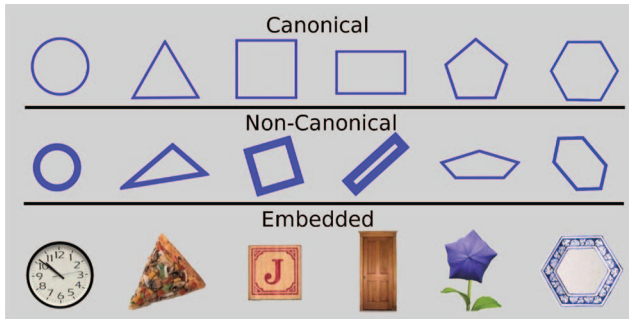


Figure 1. Stimuli used in the study. Adapted in part with permission from “The Shape of Things: The Origin of Young Children’s Knowledge of the Names and Properties of Geometric Forms,” by B. N. Verdine, K. R. Luca, R. M. Golinkoff, K. Hirsh-Pasek, and N. S. Newcombe, 2016, *Journal of Cognition and Development*, 17, p. 146. Copyright 2016 by Routledge. See the online article for the color version of this figure.

task to assess the speed with which participants looked to named target shapes, as a measure of processing.

Six categories of common geometric shapes were selected to provide a spectrum of difficulty based on performance in other studies (Fuson & Murray, 1978; Klein et al., 1999; Satlow & Newcombe, 1998; Verdine et al., 2016). Three variants of each of the six shape categories were selected to also alter difficulty across trials within the same category. The three variants included typical, monochromatic versions of shapes children were likely to have seen before, such as an equilateral triangle, referred to as *canonical*. Other versions of each shape, such as a scalene triangle, and shapes represented by everyday objects (e.g., a triangular slice of pizza), were also included and referred to as *noncanonical* and *embedded* shapes, respectively. Inclusion of these shape types was meant to challenge children’s understanding of shape names. The design also reduces the use of simple visual-matching strategies or incorrect assumptions about shapes (e.g., that shapes can only be represented by line drawings).

Hypotheses

A longitudinal design allowed us to determine whether shape processing is modulated by SES-related factors like maternal education and whether it predicts later spatial skills. We expected to find individual differences in children’s looking to the shape targets such that children with less educated mothers would take more time to locate targets and spend less time looking at them. We also expected that aspects of shape processing would predict later spatial skills.

Method

Participants

Participants were 79 typically developing 3-year-old children ($M = 3$ years; 7.7 months, $SD = 2.46$ months, range = 3 years; 1.7 months to 3 years; 11.4 months; males = 45, females = 34; higher SES = 43, lower SES = 36) which were also assessed approximately one (12.8 months; $SD = 1.6$) and two years later (25.8 months; $SD = 1.6$) at ages 4 (56.3 months; $SD = 2.8$) and 5 (69.2

months; $SD = 3.4$). Children were native monolingual English speakers from diverse preschools in a medium-sized mid-Atlantic United States city: private preschools ($N = 29$; 1 lower SES), Head Starts ($N = 30$; 1 higher SES), and a university-related preschool ($N = 19$; 6 lower and 13 higher SES).

Throughout this report we used highest level of maternal education as a proxy for SES. Hoff (2013) argued that mother’s education is a critical SES component for development. In our sample, maternal education was typically related to income; all but one child from a Head Start was characterized as lower SES and all but one child in the private preschools was characterized as higher SES. Mothers who completed a 4-year college degree or higher are referred to as the higher SES group (bachelor’s, $N = 15$; graduate, $N = 28$) and those who had less than a 4-year degree are referred to as the lower SES group (some high school, $N = 3$; high school diploma or equivalent, $N = 15$; trade qualification or 2-year associate’s degree, $N = 15$). Survey data from three children in Head Start facilities were not returned; they were characterized as lower SES.

A total of 102 children were consented. Ten children (5 higher SES; 5 lower SES) were excluded due to uncooperativeness during the eye-tracker calibration procedure or before data collection started and 13 were excluded for track loss (see the Data Analysis section). Attrition was about 20% per year of the study; 63 of the 79 participants at age 3 participated at age 4 and 50 participated at age 5. As is typical, attrition over the 2 years was higher in the lower SES group ($N = 19$ vs. 10). In Year 2, one participant was excluded for speech problems and one was excluded because the child refused to participate. All other Year 2 attrition was due to an inability to locate the participants. In Year 3 most participants matriculated to other schools and had to consent to be in the study again. All attrition was due to being unable to contact them ($N = 12$) or unwillingness to be reconsented ($N = 4$). Three participants that were not located in Year 2 were found during our Year 3 search and tested. According to a series of t tests, those that started the study but dropped out and those that completed the study performed similarly on all tests except for the Test of Spatial Assembly (TOSA); those that completed the study scored higher [$M_{\text{sample}} = .24$ ($SD = .91$)] than those that did not [$M_{\text{drops}} = -.16$ ($SD = .62$); $t(76) = 2.07$; $p = .042$]. There were SES differences in most measures (Verdine et al., 2017) and it is likely this effect resulted at least partially from differences in attrition by SES.

Apparatus and Stimuli

The SFT was conducted using a Tobii T60XL with a 1,920 × 1,200 pixel resolution monitor (75.3 × 20.8 cm) run with Tobii Studio 2.2.4 operating at a 60-Hz sampling rate. The screen was divided into left and right halves, each with a centered shape occupying approximately 9% of the screen (roughly 12 × 12 cm). The task used 18 stimuli (see Figure 1).

To select the shapes, nine variations of each shape were generated by varying vertex angles, side lengths, orientation, or line weights. The canonical representations were unanimously picked by eight adults as representing prototypical versions. They had equilateral features or a typical ratio of sizes for the sides and matched those likely encountered in books, touchscreen apps, and shape sorters (Resnick et al., 2016). Noncanonical representations were

neither typical, nor highly atypical. To select these, five versions of each shape were posted to a survey site where their typicality was rated by 73 adults on a 7-point Likert scale. The shape rating nearest the category average was selected. Highly atypical shapes were not used to avoid floor effects (Satlow & Newcombe, 1998). Certain forms such as the circle and square, by definition, had to retain properties that made them more similar to the canonical versions. Lastly, embedded shapes were photographs of everyday objects with shapes similar to the canonical forms (e.g., a rectangular door), although some were not “perfect” representations of their categories. Pizza, for example, traditionally has a rounded crust (though not in the version used) and the pentagonal flower does not have perfectly straight sides. Object selection was guided by examples in children’s books, which frequently contain imperfections. In a forced-choice paradigm, these imperfections should not greatly impact recognition if children understand even basic features of a shape.

Procedure

All testing (apart from one child who was tested in the lab) was done individually in preschools in a quiet room. The SFT in which the eye-tracker was used was administered as part of a larger longitudinal study, presented in Verdine et al. (2017). Participants sat in a chair approximately 66 cm from the monitor within the tracking range of the device (35°). After a standard 5-point calibration, children were told they were going to watch a movie and that they should look where the movie tells them, stay in their chair, and not touch the screen. Participants were presented with two practice trials and one familiarization trial (see Figure 1 in the online supplemental materials). Practice trials presented two stimuli children were expected to know (e.g., a cupcake, a tiger) and prompted them to look at one, allowing us to determine whether children understood the instructions. All participants followed these initial prompts. The familiarization trial provided a baseline of exposure to all shapes and consisted of six canonical shapes labeled in random order and presented simultaneously. Next came two blocks of nine test trials divided by a break (18 total), and four intermingled filler trials. Filler trials used the same stimuli as the practice trials and were easy trials to maintain attention (1 after Trial 5, 2 for reorientation after the break, and 1 after Trial 14).

The 18 test trials presented pairs of shapes in three quasirandom orders with the following constraints: (a) the same shape was not allowed to be the target on back-to-back trials, (b) the stimuli were always paired within representation category (e.g., only canonical shapes with canonical shapes), and (c) squares and rectangles were never paired. Each of the 18 stimuli appeared once as a target and once as a distractor. The four randomly selected prompts were recordings of either a male (for familiar object trials) or a female (for test trials) native English speaker using child-directed speech: (a) “Gosh! Check out the [target]! Check out the [target]!”; (b) “Hey! Watch the [target]! Watch the [target]!”; (c) “Oh look! Look at the [target]! Look at the [target]!”; or (d) “Wow! See the [target]! See the [target]!” The shape name prompt was repeated to reduce looking away from the target toward the end of a trial.

Each of the practice, test, and filler trials began with a 3-s video of a laughing baby or a dancing feather with a rattling sound, to center the child’s gaze. An approximately 2.5-s time came next during which the target and nontarget shapes were visible on

screen but no directions were given. Then the auditory prompt was given, and the child had approximately 3.5 s to look at the display before the fixation image for the next trial. Actual trials were 6 s; the length of time before and after a prompt varied slightly depending on the prompt (2,042–2,455 ms before and 3,535–3,948 ms after). All timing reported is relative to the onset of the first shape name (i.e., “[target]”) within the prompts. For analyses, only eye-tracker samples from 4,000 ms of the trials were used—the 1,000 ms period *before* the shape prompt (saliency phase) and the 3,000 ms period *after* onset of the prompt word (test phase). This period provides enough data to account for baseline looking prior to the prompt and a reasonable amount of time to respond. The full procedure was approximately 4 min.

Dependent variables. Areas of interest (AOIs) were the two halves of the screen with the target and the distractor. Most data are reported as the percent of eye-tracker samples (sampling rate = 60 Hz) children fixated on the target AOI versus the nontarget AOI. We refer to the amount of target looking as percent looking to target (PLT; i.e., $[\text{samples on target AOI}/(\text{samples on target AOI} + \text{samples on nontarget AOI})] \times 100$). For most analyses, PLT was calculated for each of 16 windows 250 ms in length: SaliencyWin-1 through SaliencyWin-4 in the saliency phase and TestWin-1 through TestWin-12 in the test phase. Test windows of 250 ms were selected to allow comparison of the SES groups at smaller time intervals than just the overall trial. The 250-ms test windows we selected allowed us to take some advantage of the rich time course information in our data while avoiding the statistical and reporting problems associated with the number of comparisons necessary to use smaller time units.

Note that average PLT is consistently computed across trials and we expected, a priori, that speed in locating the targets would vary on different trials. For an individual participant or group, that variance created a response curve similar to that seen in Figure 2, computed from underlying data that are inherently binary. Whether an individual child actually knows the right answer cannot be taken for granted, let alone that they will quickly arrive at that answer and consistently fixate the target without glancing, perhaps multiple times, to the distractor. Because responses to some trials take more time or are more inconsistent at first, PLT averaged across trials takes time to rise above chance and achieve a maximum. That maximum is inherently limited by the number of trials on which children know the answer and consistently fixate the target. Thinking about the answer before responding, visually scanning back-and-forth to determine the answer, simply not knowing the answer in the provided timeframe (i.e., chance looking), or being systematically wrong (i.e., PLT significantly less than .5) will all lower the average PLT for a given time window even when children are accurately responding to other trials. This is how it is possible that PLT can remain well below 100% and yet still reflect valid responding. This is also why any discussion of a “response” to the target should be assumed to be referring to the point at which PLT exceeds chance based on an average across different trials; visual behavior in any individual trial could vary significantly from the average. When this variable is included prior to processing speed in regressions, it may help to control for factors that would otherwise influence processing speed but are not directly related to it (e.g., shape familiarity, overall interest in shapes, attention to the task, etc.).

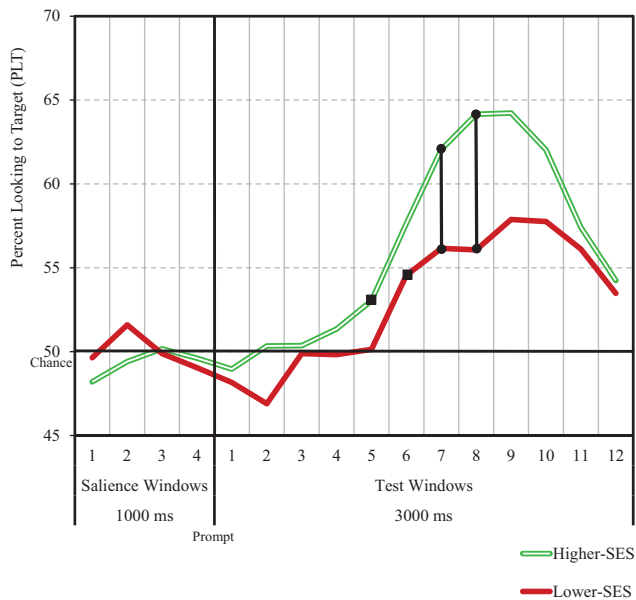


Figure 2. Percent looking to target (PLT) by socioeconomic status (SES) group on the Shape Fixation Task across all trials for each 250 ms salience and test window. The black vertical bars with circles show test windows in which the SES groups are significantly different from one another. The black square markers note the test window in which each group's PLT first statistically exceeds chance (50%). See the online article for the color version of this figure.

Longitudinal assessments. Tests of spatial skills and vocabulary were administered in the same session as the SFT at age 3 and at ages 4 and 5. All tests were untimed and did not rely on knowledge of shapes or any geometric knowledge. Though some of these tests may differentially tap certain spatial skill domains, they were combined into z-scored variables that take advantage of multiple measurements of the same construct to create more stable measurements of omnibus spatial skills. This approach helps address one of the challenges of assessing young participants, whose scores on any one test are more likely than older children and adults to be influenced by factors unrelated to the tested construct (e.g., the order of tests within a session).

Age 3 assessments. The Peabody Picture Vocabulary Test (4th ed.; PPVT; Dunn & Dunn, 2007), a measure of receptive vocabulary and also related to general intelligence (Hodapp & Gerken, 1999), asks children to point to the depiction of a prompted word from four possible images. Thus, the PPVT was used as a covariate in a number of analyses because it is a strong control for early language skills and likely also controls for some other domain-general skills that may influence performance on both tests but are not directly related to their content (e.g., inhibition control).

The TOSA was given as an early, omnibus test of spatial skills (Verdine et al., 2017). Designed for 3-year-old children, the TOSA is composed of 2- and 3-dimensional (2-D and 3-D) trials (6 of each; see Figure 2 in the online supplemental materials). Trials use two to four pieces (flat magnetic foam geometric shapes or interlocking construction blocks) which children arranged to match an always visible target design. Scores for the 2-D and 3-D trials were

calculated separately, then z scored, and those scores were averaged to create the overall TOSA scores. This measure has good reported reliability (Cronbach's alpha = .747) and validity, correlating with other established or standardized spatial measures ($r = .44$ to $.60$; Verdine et al., 2017).

Age 4 assessments. Overall age 4 spatial scores (Z age 4) were calculated by averaging the z scores from the two tests given at age 4: (a) the Woodcock-Johnson III: Spatial Relations subtest (WJ-III; Woodcock, McGrew, & Mather, 2001), and (b) the Wechsler Preschool and Primary Scale of Intelligence III: Block Design subtest (WPPSI; Wechsler, 2002). The WJ-III taps understanding of spatial relations by showing a design on the left side of the page with possible component pieces to the right. A point was awarded for each correctly identified component piece (33 trials; 81 possible points). The WPPSI Block Design subtest was administered as per the standardized instructions. Items require participants to use 3-D blocks with red and white colored patterns to create specific target designs from 2-D images.

Age 5 assessments. Overall age 5 spatial scores (Z age 5) were calculated by averaging the z scores from the two tests given at age 5: (a) the WJ-III, and (b) the Children's Mental Transformation Task (CMTT; Levine, Huttenlocher, Taylor, & Langrock, 1999; Levine, Ratliff, Huttenlocher, & Cannon, 2012). The WJ-III was the same at ages 4 and 5. The CMTT is a 32-item test that measures mental rotation and translation by asking children to select an image from among four choices that together comprise a target design.

Track Loss

Track loss occurs when the eye-tracker fails to register the child's eyes, often because the child looks away or moves. A total of 271 trials out of 1,656 (16.4%) administered to the initial 92 participants had more than 50% track loss and were removed. Individuals with seven or more excluded trials were removed from analysis ($N = 11$) and two participants with six dropped trials were excluded for missing more than three trials from a single representation type.

Results

Group-Level Analysis of Eye Gaze Patterns

Figure 2 shows the time course of looking to the target as a percent of gazes directed at the target versus the nontarget for each SES group. As the figure portrays, and as expected, during the salience windows preceding the prompt, both higher and lower SES children were near chance (50%). However, as the test windows proceed, both groups begin to rise above chance and reach a maximum PLT before looking to the target begins to decline for both groups.

To investigate group-level differences in these patterns, the group-level analyses probed for effects of SES and shape type on the timing and magnitude of responses to the shape prompts separately for each of the 16 analysis windows (4 salience and 12 test). Elogit transformations were used to reduce the dependency caused by eye gaze location not being independent from one sample to the next (Barr, 2008). Trends were tested using multi-level mixed logit modeling with crossed random intercepts for

participants and items (after Baayen, Davidson, & Bates, 2008). Due to concerns over attention, participants were given one trial for each stimulus. This trial limitation prompted two analyses to assess the magnitude of the group-level response: one with shape category (i.e., rows in Figure 1) as a fixed factor and representation type (i.e., columns in Figure 1) as a random effect, and another with representation type as a fixed factor and shape category as a random effect.¹ Preliminary analyses showed no sex differences, so that variable was excluded from further analyses. As reported in Table 13 of Verdine et al. (2017), the spatial measures given at ages 3, 4, and 5 all demonstrate significant, medium-to-large effect size SES differences (Hedges $g = .60$ to 1.69), so group-level analyses focus exclusively on SFT data in this report.

Group PLT for shape and representation categories. In the test for effects of shape category, participant and representation type were entered as crossed random intercepts and shape and SES were first-level predictors. When not perfectly correlated with random intercepts, shape was entered as a random slope in the participant effects structure, and shape, SES, and an SES \times Shape interaction were entered as random slopes in the representation type effects structure. Higher SES children had higher PLT (Table 1) in TestWin-7 (1,500–1,750 ms; $p < .05$) and TestWin-8 (1,750–2,000 ms; $p < .01$). These significant comparisons are shown in Figure 2 with vertical bars.

There was a main effect of shape for SaliencyWin-1 (–1,000 to –750 ms; $p < .01$) and effects of shape during TestWin-4 (750–1,000 ms; $p < .05$), TestWin-6 (1,250–1,500 ms; $p < .01$), and TestWin-7 (1,500–1,750 ms; $p < .05$; the best-fitting model for TestWin-7 included effects of SES and shape). There were no other effects of SES or shape and no interactions. Analysis of the specific shape effects are presented in the online-only supplemental materials.

A similar multilevel mixed model to that above tested for effects of representation type. SES effects mirrored the previous analysis, appearing in TestWin 7 and 8. The only significant effect of representation type was in TestWin 12 and appears to result from

eye gaze wandering away from canonical targets near the end of a trial, likely due to stimulus familiarity.

Group reaction time (RT): Comparison of eye gaze patterns to chance performance by SES group. To assess group-level RT to targets, we used “correct” trials only. A trial was considered correct if more than 55.56% of the eye-tracker samples in that 1,500 ms timeframe were directed toward the target. That timeframe was selected because it allowed adequate response time, even for slow responders, encompassed the peak PLT for both SES groups, and excluded the end of the trial when some children may start wandering away from the target. Assuming that a child was paying attention to the prompt, knew the name of the shape, and was responding appropriately to a trial, their looking to the target in this timeframe should be above chance. The 55.56% chance threshold was calculated by determining the lowest k value with a probability (i.e., $P = p$ value) below 0.05. $P(k \text{ successes in } n \text{ samples}) = [n!/k!(n - k)!]p^k q^{n-k}$, where $n =$ eye-tracker samples in the time window (i.e., 60 samples/s \times 1.5 s = 90); $k =$ on-target samples; $p =$ probability of success (chance = 0.5); and $q = 1 - p =$ probability of failure. P drops below 0.05 with 50 on-target samples out of 90 (55.56%; actual probability = 0.048).

Speed analyses for the group and individual analyses were focused on these correct trials to remove the influence of trials on which children were inattentive or for some reason did not look very much at the target shape (i.e., low PLT trials). This was a valid concern since the higher SES group had a higher proportion of correct trials than the lower SES group for noncanonical targets ($M_{\text{high}} = .62$ ($SD = .21$); $M_{\text{low}} = .51$ ($SD = .24$); $t(77) = 2.25$; $p = .027$) and squares ($M_{\text{high}} = .69$ ($SD = .35$); $M_{\text{low}} = .52$ ($SD = .31$); $t(77) = 2.16$; $p = .034$), with a trend toward more correct overall ($M_{\text{high}} = .62$ ($SD = .16$); $M_{\text{low}} = .55$ ($SD = .15$); $t(77) = 1.89$; $p = .062$). Practically, it would be odd to measure the speed of responses that did not appear to actually identify the target.

PLT {PLT = [target looks/(target looks + nontarget looks)] \times 100} in the sixteen 250- ms windows was compared with chance by assessing the t values of the intercepts of empty models with no fixed factors. Elogit-transformed PLT was modeled for each SES group for each window using multilevel mixed elogit modeling with participant and representation type as crossed random intercepts. Higher SES children were faster than lower SES children, first demonstrating above chance PLT in TestWin-5 (1,000–1,250 ms; $t = 2.39$, $p < .05$) rather than TestWin-6 (1,250–1,500 ms; $t = 4.90$, $p < .05$). Therefore, the higher SES group looked more to the targets and were faster in shifting attention to them. The SES difference in speed is unlikely due to differences in the groups’ ability to look to the correct targets since only correct trials were used.

Table 1
Statistically Significant Effects of SES From the Best-Fitting Multilevel Linear Models of PLT by Analysis Window

| Effect | Estimate | SE | t |
|---|----------|-----|-------|
| TestWin-7 (1,500–1,750 ms after prompt) | | | |
| Intercept | 1.22 | .21 | 5.95 |
| Higher-SES vs. lower-SES | –.41 | .17 | –2.38 |
| TestWin-8 (1,750–2,000 ms after prompt) | | | |
| Intercept | .98 | .13 | 7.52 |
| Higher-SES vs. lower-SES | –.55 | .19 | –2.87 |

Note. SES = socioeconomic status; PLT = Percent looking to target. The models presented are the best-fitting model for each time window; when effects or interactions do not appear, it is because adding them to the models did not reliably improve the fit. Formula in R: TestWin-7: DepVar \sim SES + Shape + (1 + Shape | Subject) + (1 | Rep Type); TestWin-8: DepVar \sim SES + (1 | Subject) + (1 | Rep Type). For the remaining 14 analysis windows, no models tested provided a better fit than an empty model with no fixed effects.

¹ The t values reported in Table 1 and throughout the group-level shape category section are for testing aspects of the multilevel hierarchical logit models used for these group-level analyses. They are not t values from traditional t tests or analyses of variance. The lmer/lme4 package available for the R statistical program does not report degrees of freedom or p values for these t statistics because these are not appropriate calculations for linear mixed models (Bates, 2006), and their omission was intentional. For comparison, t statistics with absolute values greater than approximately 2.0 would have p values less than 0.05.

Individual-Level Analyses of Shape Processing

Individual PLT and speed variables. A PLT score for each child was calculated by determining average PLT across all eye-tracker samples in all test windows for all trials. Creating the speed variable for each individual required identifying how long each child took, on average, to respond to targets for correct trials. This was operationalized by identifying the window in which each child, on average, tended to move their gaze most toward the target compared with the previous test window. An average PLT was calculated across all correct trials for each of the twelve 250-ms test phase windows for each participant. The consecutive window pair with the largest positive slope between them was selected as the average “response” to the target for each child (i.e., the pair where PLT for the second window minus PLT for the first window was highest). Once selected, the number for the second test window of the pair was assigned as the speed of the child’s response, representing the time where the participant moved most from off- or nontarget looking to looking at the target. TestWin-1 and 2 (0–500 ms) were not assigned speed values as Fernald and Marchman (2012) indicated that looks that fast do not reflect reactions to the prompt. Thus, speed values were whole numbers from 3 to 12; those having lower numbers were faster responders.²

Individual-level analyses using PLT and speed as predictors. PPVT scores were correlated with PLT ($r = .26, p = .022; df = 76$). PPVT scores were used in individual-level analyses as a covariate for receptive vocabulary and domain-general skill. Participant age was not significantly correlated with scores on any of the measures except the TOSA ($r = .40, p < .001; df = 76$). However, since the age range of the sample was approximately one year, age was entered as a covariate.

The speed variable used only correct trials, in part to allow speed and PLT to vary independently, and indeed they were not correlated ($r = -.12; ns$) even after partialing out PPVT scores ($pr = -11; ns$). Correlations between the SFT variables and all of the other tests are reported in Table 2. Scatterplots showing the PLT and speed variables from the SFT with z -score combined spatial scores at age 5 appear in Figure 3. Table 2 shows that PLT at age 3 was related to performance on the age 3 spatial measure (the TOSA), even after accounting for PPVT scores and age. PLT was not related to WJ-III scores at age 4 or age 5, but PLT was marginally related to WPPSI scores at age 4 and significantly related to CMTT scores at age 5. With respect to speed, our data hint that the relation between speed at age 3 and spatial skills gets stronger over time, speed was not related to spatial ability at age 3 ($r = -.06$), there were mixed results at age 4 ($r = -.22$ to $-.29$), and there were significant correlations with both spatial measures and the z -scored spatial measure at age 5 ($r = -.33$ to $-.41$). According to Fisher r -to- z transformations, the correlations between speed and the spatial measure at age 3 (TOSA) and speed and the age 5 z -score combined spatial measure are significantly different ($z = -1.98; p = .048$). Though the correlations between speed with spatial skill and other variables from age 3 to age 4 and age 4 to age 5 were not significant, there was a consistent pattern: the apparent magnitude of longer-term correlations were higher. Even the correlations of speed with the WJ-III follow this pattern ($r_{Age4} = -.22; r_{Age5} = -.39$). As it is unusual for correlations over increasingly large time frames to appear to increase, this pattern might be explored in future work.

To understand how PLT and speed are related to spatial ability after accounting for other covariates and earlier spatial skills, regressions were run predicting age 5 spatial skills from the other variables in the study (Table 3). Our approach to presenting these regressions was to start with the simplest model using only our PLT and speed variables, to understand how well they would predict later spatial skills if the SFT was the only measure given. We then systematically added groups of variables to control for and understand their influence. The second model adds individual difference variables that we thought were likely to influence performance broadly and represents a parsimonious and reasonably controlled model for predicting later spatial skills. The third model controlled for age 3 and 4 spatial skills before adding the PLT and speed variables. We did this to see if the PLT and speed variables are just proxy measures of spatial skill. If true, neither variable should be significant in Model 3 even if they are strong predictors in Models 1 and 2.

When entered into a regression without covariates, PLT and speed account for 26.3% of the variability in spatial skills 2 years later at age 5 (Model 1). The two variables have similar standardized beta weights in the model (PLT = .353; speed = $-.372$) and account for similar variability in age 5 spatial skills (Adj. R^2 PLT = .139; speed = .124). As suggested by the low correlation of PLT with speed, order of entry had negligible effects and we present all of the models entering PLT first. Both variables showed important and independent contributions to the prediction of later spatial skills when no control variables were used.

Model 2 entered SES, PPVT scores, and age in Step 1 to account for individual differences related to parent education, participant vocabulary, and age-related maturation. These variables accounted for significant variability in age 5 spatial scores (Adj. $R^2 = .412$), but both PLT and speed remained significant predictors in Model 2 (Δ Adj. $R^2 = .048$ and $.102$, respectively). Model 3 added scores from the TOSA in Step 2 and the z -score combined average of the WJ-III and WPPSI in Step 3, to account for individual differences in spatial skills at ages 3 and 4. The overall model accounted for over 81% of the variability in age 5 spatial skills (Adj. $R^2 = .815$). Notably, the inclusion of either the TOSA or the age 4 spatial measures caused PLT to lose prediction potency. This indicates a significant overlap in the amount of variability in age 5 spatial skills explained by earlier spatial skills and PLT. This finding comports with our hypothesis that identifying named shapes is related to spatial development. However, speed remained a significant predictor in this model, accounting for about 5.8% of the variability in spatial skills assessed 2 years later. Thus, speed in

² One caveat for assigning the speed value is that the percent looking to target (PLT) in the second window of the pair had to be above 55.56%. This rule ensured that the child fixated the target long enough on average to be considered “on target” after the response. If the second window was not above 55.56% PLT for the pair with the largest slope, the pair with the next largest slope was considered until a pair was found which had the largest slope possible and PLT higher than 55.56% for the second window. Of the nine participants with data that violated this rule, for eight of them the pair with the second largest slope met the criterion. The cutoff of 55.56% was chosen from our calculation of chance for the group RT analyses, but is not truly reflective of chance for all children because their number of correct trials varied. Instead of changing the chance cutoff by participant, a conservative chance estimate (based on $n = 90$) was chosen and applied to all participants.

Table 2
Correlations and Partial Correlations Controlling for PPVT and Age

| Age | <i>df</i> ^a | Test type | Variable | Age 3 | | | | Age 4 | | | Age 5 | | | Age ^c |
|---|------------------------|------------|----------------------|------------|--------|----------------|--------------|----------------|--------|-------|----------------|--------|--------|------------------|
| | | | | Shape task | | Spatial (TOSA) | Vocab (PPVT) | Spatial | | | Spatial | | | |
| | | | | Acc | Speed | | | Z ^b | WJ-III | WPPSI | Z ^b | WJ-III | CMTT | |
| Zero-order Pearson correlations | | | | | | | | | | | | | | |
| 3 | 76 | Shape task | PLT | | -.12 | .36** | .26* | .24 | .10 | .32 | .40** | .23 | .46** | -.02 |
| | | | Speed | -.12 | | -.06 | -.02 | -.29* | -.22 | -.29* | -.41** | -.39** | -.33* | -.01 |
| | | Spatial | TOSA | .36** | -.06 | | .40** | .58** | .46** | .54** | .68** | .58** | .60** | .40** |
| | | Vocab | PPVT | .26* | -.02 | .40** | | .48** | .41** | .41** | .59** | .46** | .56** | .19 |
| 4 | 61 | Spatial | Z ^b | .24 | -.29* | .58** | .48** | | .86** | .87** | .80** | .67** | .74** | .09 |
| | | | WJ-III | .10 | -.22 | .46** | .41** | .86** | | .48** | .71** | .62** | .63** | .05 |
| | | | WPPSI | .32* | -.29* | .54** | .41** | .87** | .48** | | .69** | .55** | .66** | .10 |
| 5 | 48 | Spatial | Z ^b | .40** | -.41** | .68** | .59** | .80** | .71** | .69** | | .87** | .88** | .08 |
| | | | WJ-III | .23 | -.39** | .58** | .46** | .67** | .62** | .55** | .87** | | .52** | .12 |
| | | | CMTT | .46** | -.33* | .60** | .56** | .74** | .63** | .66** | .88** | .52** | | .02 |
| | | | Age ^c | -.02 | -.01 | .40** | .19 | .09 | .05 | .10 | .08 | .12 | .02 | |
| Partial correlations controlling for PPVT and age at time of test | | | | | | | | | | | | | | |
| 3 | 74 | Shape task | PLT | | -.11 | .34** | | .14 | -.01 | .24 | .31* | .13 | .39** | |
| | | | Speed | -.11 | | -.06 | | -.32* | -.23 | -.31* | -.49** | -.43** | -.38** | |
| | | Spatial | TOSA | .34** | -.06 | | | .52** | .40** | .47** | .66** | .51** | .58** | |
| 4 | 59 | Spatial | Z age 4 ^a | .14 | -.32* | .52** | | | .82** | .84** | .73** | .57** | .66** | |
| | | | WJ-III | -.01 | -.23 | .40** | | .82** | | .38** | .63** | .53** | .53** | |
| | | | WPPSI | .24 | -.31* | .47** | | .84** | .38** | | .61** | .44** | .58** | |
| 5 | 46 | Spatial | Z age 5 ^a | .31* | -.49** | .66** | | .73** | .63** | .61** | | .83** | .82** | |
| | | | WJ-III | .13 | -.43** | .51** | | .57** | .53** | .44** | | | .36* | |
| | | | CMTT | .39** | -.38** | .58** | | .66** | .53** | .58** | .82** | .36* | | |

Note. TOSA = Test of Spatial Assembly; PPVT = Peabody Picture Vocabulary Test; WJ-III = Woodcock-Johnson III; WPPSI = Wechsler Preschool and Primary Scale of Intelligence III; CMTT = Children's Mental Transformation Task; PLT = percent looking to target.

^a Degrees of freedom apply pairwise and are limited by the variable with the lowest degrees of freedom (e.g., all zero-order CMTT correlations *df* = 48). ^b The Z variables are summary variables averaging z scores from the two individual spatial tests listed below them (WJ-III and WPPSI at age 4; WJ-III and CMTT at age 5). The score provides a stronger measure of the construct by collapsing across multiple theoretically-related tests within a time point. ^c Age = age in months at the time of the age 3 testing session.

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

processing shape names is not only an important predictor of spatial skills but has a unique influence not explained by receptive vocabulary or spatial skills measured with untimed tests, or even the amount of looking to target shapes within the same task.

Discussion

A recent emphasis on teaching early geometry in preschool and kindergarten (Common Core State Standards Initiative, 2010; Office of Head Start, 2011) and evidence suggesting the importance of early spatial skills in long-term STEM outcomes (e.g., Wai et al., 2010) prompted this study. Young children's speed of processing familiar, everyday words (e.g., Hurtado et al., 2008) inspired us to probe the speed with which preschoolers identify named geometric shapes. There are some differences in response to different shape's types but there were no effects of representation type, which we expected especially for noncanonical shapes. It seems likely that children were able to bootstrap their knowledge of canonical shapes to identify them (e.g., the noncanonical triangle still looks more like a triangle than the noncanonical rectangle), whether or not they really understood the *defining* features of the shapes (Satlow & Newcombe, 1998). Aside from these findings, this study was primarily aimed at exploring whether differences in maternal education influence shape identification and how individual differences in identifying shapes predict later spatial skills.

Does SES Influence Shape Identification?

The present data confirm that even when language is from a specific content area appearing with relatively low frequency in speech (i.e., geometric shape terms; Rudd et al., 2008; Sarama & DiBiase, 2004), higher SES children have an advantage. They move more quickly to targets than the lower SES group, rising above chance looking to target in Test Window 5 rather than Test Window 6 (approximately 250 ms), indicating that they are processing the language more quickly. They also look more at target shapes than the lower SES group, on average, in Test Windows 7 and 8.

A long history of research suggests that SES factors influence the amount and type of language children hear (e.g., Hart & Risley, 1995) and that these experiences influence vocabulary size and processing speed (Weisleder & Fernald, 2013; Hurtado et al., 2008). Hurtado and colleagues (2008) showed that the amount of language low-income parents direct to their children influences their children's processing speed at 24 months. For children whose parents used more words during a play session, they found an approximately 200-ms advantage in children's RT to familiar nouns. Analogously, we found that participants with more educated mothers, who we know tend to hear more language (Hart & Risley, 2003; Hoff, 2013), respond one 250-ms window faster than

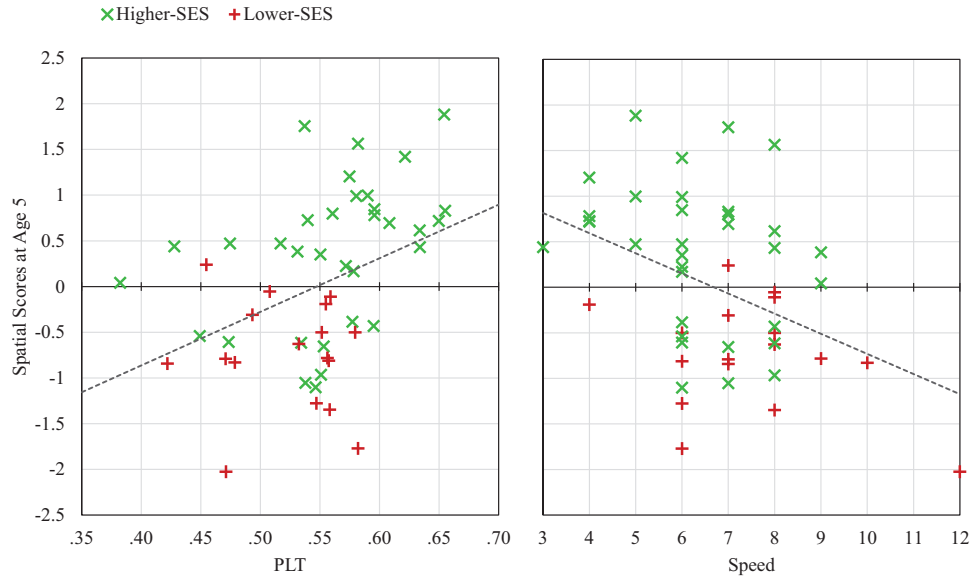


Figure 3. Scatterplots with linear trend lines showing the relation between spatial scores at age 5 and the Shape Fixation Task variables (percent looking to target and speed). See the online article for the color version of this figure.

peers with less educated mothers. Thus, the speed difference potentially lends additional credence to those earlier findings.

It appears then that lower SES children are at a disadvantage in processing the names of geometric shapes compared with their higher SES peers. Assuming these differences remain relatively constant and that shape identification predicts later spatial skills, this may have the unfortunate consequence of impeding STEM learning for lower SES children.

Do Individual Differences in Shape Processing Predict Future Spatial Skills?

A delay of approximately 250 ms in processing and responding to prompted shape words at age 3 might seem unlikely to have practical, long-term impacts on learning. However, research has documented long-term impacts of similar sized delays in the recognition of everyday words (e.g., Fernald & Marchman, 2012;

Table 3
Results of Hierarchical Multiple Regressions for Predicting Spatial Scores^a at Age 5

| Model | Step | Variable | ΔF^b | df | Adj. R^2 | Δ Adj. R^2 | Unstandardized coefficients | | Standardized coefficients | | | Partial r |
|-------|------|----------------------|--------------|-------|------------|---------------------|-----------------------------|------|---------------------------|-------|------|-------------|
| | | | | | | | B | SE | β | t | p | |
| 1 | 1 | Constant | | | | | -1.49 | 1.11 | | -1.34 | .186 | |
| | 2 | PLT | 8.91** | 1, 48 | .139 | .139 | 5.13 | 1.80 | .353 | 2.86 | .006 | .385 |
| 2 | 1 | Constant | | | | | -2.20 | .07 | | -3.01 | .004 | -.402 |
| | 2 | Speed | 9.74** | 2, 47 | .263 | .124 | -.20 | .07 | -.372 | -3.01 | .004 | -.402 |
| 3 | 1 | Constant | | | | | -3.35 | 1.90 | | -1.77 | .084 | |
| | 1 | SES | 12.46*** | 3, 46 | .412 | .412 | .57 | .22 | .315 | 2.61 | .012 | .366 |
| | | PPVT | | | | | .02 | .01 | .340 | 2.77 | .008 | .385 |
| | | Age | | | | | .02 | .04 | .060 | .60 | .554 | .089 |
| | 2 | PLT | 5.03* | 4, 45 | .460 | .048 | 3.04 | 1.44 | .060 | 2.12 | .040 | .304 |
| | 3 | Speed | 11.47** | 5, 44 | .562 | .102 | .05 | .05 | -.328 | -3.39 | .002 | -.412 |
| 3 | 1 | Constant | | | | | 1.21 | 1.56 | | .78 | .442 | |
| | 1 | SES | 11.65*** | 3, 43 | .410 | .410 | .21 | .16 | .118 | 1.38 | .176 | .215 |
| | | PPVT | | | | | .01 | .00 | .207 | 2.40 | .021 | .359 |
| | | Age | | | | | -.04 | .03 | -.118 | -1.50 | .141 | -.234 |
| | 2 | TOSA | 26.05*** | 4, 42 | .627 | .217 | .38 | .10 | .352 | 3.71 | .001 | .511 |
| | 3 | Z age 4 ^a | 22.44 | 5, 41 | .753 | .126 | .37 | .09 | .363 | 4.01 | .000 | .541 |
| | 4 | PLT | 1.73 | 6, 40 | .757 | .004 | 1.05 | 1.03 | .072 | 1.02 | .315 | .161 |
| | 5 | Speed | 13.33*** | 7, 39 | .815 | .058 | -.13 | .04 | -.252 | -3.65 | .001 | -.505 |

Note. PLT = percent looking to target; SES = socioeconomic status; PPVT = Peabody Picture Vocabulary Test; TOSA = Test of Spatial Assembly. ^aThe “Z age 4” variable is a summary variable calculated by taking the average z score from the Woodcock-Johnson III and Wechsler Preschool and Primary Scale of Intelligence III given at age 4. ^bAsterisks in this column indicate the p value of the change in F for a given step of the regression. * p < .05. ** p < .01. *** p < .001.

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008). Our individual-level analyses suggest that processing speed for shape names relates to the learning of spatial skills over the ensuing two years and helps substantiate the conclusion that speed is consequential over-and-above simply identifying word targets. Speed was not related to concurrent spatial skills at age 3 but was a unique predictor of later spatial skills. Furthermore, speed as a predictor of age 5 spatial skills was affected little by the inclusion of a wide range of controls.

On the contrary, the variance explained by PLT, our measure of attention, was especially influenced by the inclusion of scores from early spatial tests. The impact of entering spatial scores into the model on the predictive power of PLT indicates that there is a substantial degree of overlap between PLT and the spatial variables. Thus, both looking to the target shapes and identifying them quickly appear to be notable independent predictors of later spatial skills. However, though PLT from our SFT has some merit as an early test of spatial skill and does predict later spatial skills, it offers little for prediction of later spatial skills above and beyond other available spatial measures. A strong overlap of this sort would be expected if shape identification relies on or helps build spatial skills and, thus, these results align with our suggestion that the SFT has a spatial component. Alternatively, speed predicts unique variability not accounted for by a host of other controls, which points to the speed variable being associated with a cause not controlled for in our models.

Many of the domain-general explanations for the link between speed and later spatial skills were controlled for by our covariates (vocabulary knowledge, SES, and age). Though we cannot rule out that the speed variable simply tapped into general language processing speed (e.g., Fernald, Marchman, & Weisleder, 2013), it is likely that including vocabulary and SES among our predictors controlled for much of that variability. Nonetheless, we are unable to determine from our correlational design whether response speed in the SFT is an index of general speed of processing, language-specific speed of processing, speed of processing shape terms specifically, or perhaps even the result of some other cognitive factor like executive function. Certainly the SFT requires processing of language other than *just* shape terms. There is a possibility that factors influencing language processing in general could play some role in quickly reacting to the SFT and later spatial tasks. For example, response speed to general nouns is related to later working memory (Marchman & Fernald, 2008) and working memory may be a factor in the tasks used here. However, aside from the controls used, the fact that language and spatial ability are usually not strongly related also makes it unlikely that general language processing skills could be the proximal cause of the observed link between speed and spatial skills. Most standardized tests treat language and spatial domains as separable and research on working memory seems to indicate largely separable systems for processing language and visual-spatial information (Shah & Miyake, 1996), except perhaps with written language, which is inherently spatial (Baddeley, 2003).

Though we have argued that it is unlikely that the speed link to later spatial skills is explained by general language speed advantages alone, it seems rather unlikely that the link we have observed is so specific that it only applies to speed in identifying a handful of geometric shapes; the link with speed may generalize in response to other spatial or number words. It is also conceivable that

differences in things like how children deploy their attention could result in slower responses and that the observed outcomes in this and other studies are not related to specific or general language processing at all. For example, a child who anticipates the prompt and has their attention focused may be faster to respond than a child whose attention wanders at the start of a trial. More work will be needed to untangle the extent to which the link between speed and later spatial skills is specific to shape terms or language. Nonetheless, this work shows significant links between shape processing, measured either with percent or speed of looking to targets, and later spatial skills.

A long-term relation of individual differences in speed at age 3 and spatial skills at age 5, but an absence of a short-term relation between speed and spatial skills at age 3 may seem surprising. However, this pattern could be expected if fluency in using geometric and spatial terms creates a *foundation* for future learning; children who process these words faster may use that skill to understand more spatial language and better build their spatial skills over the ensuing two years. For example, if a parent is describing the spatial properties of a shape but the child is processing all language more slowly including shape names, the child may fail to actually think about the spatial properties the parent is highlighting. Over time, this might be expected to slow a child's mastery of spatial language and skills. Such a mechanism is speculative and requires more study.

Differences in the length of time it takes to process spatial language, as opposed to differences in accuracy, may have important downstream implications for success in STEM-specific content areas. While children are resolving spatial or geometric language, they are using cognitive resources that might be better deployed for assimilating new knowledge (e.g., Just & Carpenter, 1980). Novel, challenging, and technical information is often encountered in STEM subjects so these deficits may have important long-term implications for STEM success.

Limitations

The preferential looking method, despite its advantages for observing the time course of responses while keeping task requirements low, has some limitations for 3-year-old children. For example, children could fixate the target briefly and then begin exploring the visual scene. In doing so, some children could move off target, despite having responded to the prompt, and lower their group's PLT. Though this is a concern in any looking time study, it is mitigated by the use of a second prompt during the trial, which redirected children to the target. Further, the effect of this is more likely to have reduced group differences that may have otherwise been observed at the end of the trial than it is to have created any of the group differences noted near the beginning; the observed SES differences appear shortly after the prompt on the upswing of each group's response.

Track loss also resulted in some children (3 higher SES; 10 lower SES) being dropped from the study and these drops do not appear to be entirely random. These participants had raw PPVT scores that were about half of a standard deviation below the analyzed sample ($M_{\text{Track Loss}} = 59.9$, $SD = 21.3$; $M_{\text{Full Sample}} = 68.1$; $SD = 21.1$) and tended to be concentrated on both ends of the range of PPVT scores. Of the dropped participants 10 had PPVT standard scores below the analyzed sample average of 109.1 ($SD = 19.9$). All of the remain-

ing three participants had high PPVT scores more than one standard deviation above the population mean (121, 127, and 142). Though dropping these children was necessary, the drops likely impact how representative the sample is of the overall population. Nonetheless, the analyzed sample was still diverse relative to most laboratory-based research.

Conclusions

Unfortunately, this study joins other literature demonstrating that 3-year-old children from lower SES backgrounds are already behind in language processing (Hart & Risley, 2003) and spatial skills (Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). We now know that processing speed differences exist in both general vocabulary (Fernald et al., 2013) and geometric terms (this study). However, this paper uniquely links processing speed in a particular domain of language (shape terms) to long-term outcomes in a related, yet quite different domain (spatial skills). As preschool skills in understanding and describing shapes claim a more important place in early education (Common Core State Standards Initiative, 2010; Office of Head Start, 2011), largely due to their short- and long-term implications for STEM success (Verdine et al., 2017; Wai, Lubinski, & Benbow, 2009), this work should be taken as a call for further study of whether these relationships are causal and how best to leverage them if they are.

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. <http://dx.doi.org/10.1016/j.jml.2007.12.005>
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208. [http://dx.doi.org/10.1016/S0021-9924\(03\)00019-4](http://dx.doi.org/10.1016/S0021-9924(03)00019-4)
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474. <http://dx.doi.org/10.1016/j.jml.2007.09.002>
- Bates, D. (2006, May 19). *Lmer, p-values and all that* (R-help mailing list). Retrieved from <https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html>
- Benjafield, J. (1976). The ‘golden rectangle’: Some new data. *The American Journal of Psychology*, *89*, 737–743. <http://dx.doi.org/10.2307/1421471>
- Clements, D. H., Swaminathan, S., Hannibal, M. A. Z., & Sarama, J. (1999). Young children’s concepts of shape. *Journal for Research in Mathematics Education*, *30*, 192–212. <http://dx.doi.org/10.2307/749610>
- Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Retrieved from <http://www.corestandards.org/Math>
- Dunn, D. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Bloomington, MN: Pearson Assessments.
- Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*, 203–222. <http://dx.doi.org/10.1111/j.1467-8624.2011.01692.x>
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234–248. <http://dx.doi.org/10.1111/desc.12019>
- Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98–116. <http://dx.doi.org/10.1037/0012-1649.42.1.98>
- Fernald, A., & Weisleder, A. (2011). Early language experience is vital to developing fluency in understanding. In S. Neuman & D. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 3, pp. 3–20). New York, NY: Guilford Press.
- Fuson, K., & Murray, C. (1978). The haptic-visual perception, construction, and drawing of geometric shapes by children aged two to five: A Piagetian extension. In R. Lesh & D. Mierkiewicz (Eds.), *Concerning the development of spatial and geometric concepts* (pp. 49–83). Columbus, OH: ERIC Clearinghouse.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children* (Vol. XXIII). Baltimore, MD: Paul H Brookes.
- Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, *27*, 4–9.
- Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., . . . Suma, K. (2015). The contribution of early communication quality to low-income children’s language success. *Psychological Science*, *26*, 1071–1083. <http://dx.doi.org/10.1177/0956797615581493>
- Hodapp, A. F., & Gerken, K. C. (1999). Correlations between scores for Peabody Picture Vocabulary Test-III and the Wechsler Intelligence Scale for Children-III. *Psychological Reports*, *84*, 1139–1142. <http://dx.doi.org/10.2466/pr0.1999.84.3c.1139>
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*, 4–14. <http://dx.doi.org/10.1037/a0027238>
- Horton-Ikard, R., & Weismer, S. E. (2007). A preliminary examination of vocabulary and word learning in African American toddlers from middle and low socioeconomic status homes. *American Journal of Speech-Language Pathology*, *16*, 381–392. [http://dx.doi.org/10.1044/1058-0360\(2007\)041](http://dx.doi.org/10.1044/1058-0360(2007)041)
- Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*, F31–F39. <http://dx.doi.org/10.1111/j.1467-7687.2008.00768.x>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354. <http://dx.doi.org/10.1037/0033-295X.87.4.329>
- Klein, A., Starkey, P., & Wakeley, A. (1999, April). *Enhancing pre-kindergarten children’s readiness for school mathematics*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada. Retrieved from <http://eric.ed.gov/?id=ED431556>
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, 299–321. [http://dx.doi.org/10.1016/0885-2014\(88\)90014-7](http://dx.doi.org/10.1016/0885-2014(88)90014-7)
- Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*, 940–949. <http://dx.doi.org/10.1037/0012-1649.35.4.940>
- Levine, S. C., Ratliff, K. R., Huttenlocher, J., & Cannon, J. (2012). Early puzzle play: A predictor of preschoolers’ spatial transformation skill. *Developmental Psychology*, *48*, 530–542. <http://dx.doi.org/10.1037/a0025913>
- Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science*, *16*, 841–845. <http://dx.doi.org/10.1111/j.1467-9280.2005.01623.x>
- Magnuson, K. A., Meyers, M. K., Ruhm, C. J., & Waldfogel, J. (2004). Inequality in preschool education and school readiness. *American Educational Research Journal*, *41*, 115–157. <http://dx.doi.org/10.3102/00028312041001115>

- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11*, F9–F16. <http://dx.doi.org/10.1111/j.1467-7687.2008.00671.x>
- Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math: Developmental and educational implications. In J. B. Benson (Ed.), *Advances in child development and behavior* (Vol. 42, pp. 197–243). Waltham, MA: Academic Press. <http://dx.doi.org/10.1016/B978-0-12-394388-0.00006-X>
- Office of Head Start. (2011). *The Head Start child development and early learning framework: Promoting positive outcomes in early childhood programs serving children 3–5 years old*. Washington, DC: U.S. Department of Health and Human Services. Retrieved from [http://eclkc.ohs.acf.hhs.gov/hslc/ta-system/teaching/eecd/Assessment/Child%20Outcomes/HS_Revised_Child_Outcomes_Framework\(rev-Sept2011\).pdf](http://eclkc.ohs.acf.hhs.gov/hslc/ta-system/teaching/eecd/Assessment/Child%20Outcomes/HS_Revised_Child_Outcomes_Framework(rev-Sept2011).pdf)
- Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science, 14*, 1417–1430. <http://dx.doi.org/10.1111/j.1467-7687.2011.01088.x>
- Rakison, D. H., & Oakes, L. M. (2003). *Early category and concept development: Making sense of the blooming, buzzing confusion*. New York, NY: Oxford University Press.
- Resnick, I., Verdine, B. N., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Geometric toys in the attic? A corpus analysis of early exposure to geometric shapes. *Early Childhood Research Quarterly, 36*, 358–365. <http://dx.doi.org/10.1016/j.ecresq.2016.01.007>
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology, 4*, 328–350. [http://dx.doi.org/10.1016/0010-0285\(73\)90017-0](http://dx.doi.org/10.1016/0010-0285(73)90017-0)
- Rudd, L. C., Lambert, M. C., Satterwhite, M., & Zaier, A. (2008). Mathematical language in early childhood settings: What really counts? *Early Childhood Education Journal, 36*, 75–80. <http://dx.doi.org/10.1007/s10643-008-0246-3>
- Sarama, J., & Clements, D. H. (2004). Building Blocks for early childhood mathematics. *Early Childhood Research Quarterly, 19*, 181–189. <http://dx.doi.org/10.1016/j.ecresq.2004.01.014>
- Sarama, J., & DiBiase, A. M. (2004). The professional development challenge in preschool mathematics. In D. H. Clements & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 415–448). Mahwah, NJ: Erlbaum.
- Satlow, E., & Newcombe, N. S. (1998). When is a triangle not a triangle? Young children's developing concepts of geometric shape. *Cognitive Development, 13*, 547–559. [http://dx.doi.org/10.1016/S0885-2014\(98\)90006-5](http://dx.doi.org/10.1016/S0885-2014(98)90006-5)
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4–27. <http://dx.doi.org/10.1037/0096-3445.125.1.4>
- Shutts, K., Örnkloo, H., von Hofsten, C., Keen, R., & Spelke, E. S. (2009). Young children's representations of spatial and functional relations between objects. *Child Development, 80*, 1612–1627. <http://dx.doi.org/10.1111/j.1467-8624.2009.01357.x>
- Smith, L. B. (2003). Learning to recognize objects. *Psychological Science, 14*, 244–250. <http://dx.doi.org/10.1111/1467-9280.03439>
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*, 352–402. <http://dx.doi.org/10.1037/a0028446>
- Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2017). Links between spatial and mathematical skills across the preschool years. *Monographs of the Society for Research in Child Development, 82*, 1–150. <http://dx.doi.org/10.1111/mono.12263>
- Verdine, B. N., Lucca, K. R., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2016). The shape of things: The origin of young children's knowledge of the names and properties of geometric forms. *Journal of Cognition and Development, 17*, 142–161. <http://dx.doi.org/10.1080/15248372.2015.1016610>
- 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*, 817–835. <http://dx.doi.org/10.1037/a0016127>
- 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*, 860–871. <http://dx.doi.org/10.1037/a0019454>
- Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence* (3rd ed.). Sydney, New South Wales, Australia: Pearson.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*, 2143–2152. <http://dx.doi.org/10.1177/0956797613488145>
- Woodcock, R. W., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson Tests of Achievement*. Itasca, IL: Riverside.

Received May 13, 2016

Revision received May 1, 2017

Accepted May 16, 2017 ■