Early Childhood Research Quarterly 46 (2019) 126–141

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

Effects of geometric toy design on parent–child interactions and spatial language

Brian N. Verdine[a,*], Laura Zimmermann[a], Lindsey Foster[a], Maya A. Marzouk[b], Roberta Michnick Golinkoff[a], Kathy Hirsh-Pasek[b], Nora Newcombe[c]

[a] University of Delaware, United States
[b] Yeshiva University, United States
[c] Temple University, United States

ARTICLE INFO

Article history:
Received 24 May 2017
Received in revised form 23 February 2018
Accepted 29 March 2018
Available online 18 May 2018

Keywords:
Language
Geometry
Spatial
Touchscreen
Preschool
Digital

ABSTRACT

Geometric forms have formal definitions. While knowing shape names is considered important for school-readiness, many children do not understand the defining features of shapes until well into elementary school (Satlow & Newcombe, 1998). One reason is likely that they do not encounter enough variety in the shapes they see (Resnick, Verdine, Golinkoff, & Hirsh-Pasek, 2016). The present study observed 60 parents and their 3-year-old children during play with geometric toys, exploring how spatial language varied with the nature of the shape-toy set (canonical shapes versus a mix of canonical and unusual or less-canonical variants) and whether geometric shapes were presented as tangible, traditional toys or shown on a touchscreen tablet app. Although children in the app condition heard more shape names than the other conditions due to the language produced by the app itself, children used more overall words and more spatial language with tangible toys that included varied shapes. In addition, parents used more shape names with sons than with daughters and tended to adjust their use of spatial language more in response to varied shape sets with boys, although these findings need replication to evaluate generality. These data suggest that including non-canonical shapes in tangible shape toys may provide a low-cost, high-impact way of refining adult-child interactions that might facilitate children's early geometric knowledge.

Shape knowledge is important for school-readiness and is an integral part of the early childhood standards for beginning math (Common Core State Standards Initiative, 2010). Accordingly, shapes have also become a focus of preschool curricula (Office of Head Start, 2011). Shape knowledge represents an early form of geometric knowledge, which is increasingly included in broadly focused early math measures, such as the Child Math Assessment (Starkey, Klein, & Wakeley, 2004) or the Research-Based Early Maths Assessment (Clements, Sarama, & Liu, 2008). Shape knowledge is, thus, widely regarded as a vital part of a foundation for future mathematical development

Research also indicates that identifying geometric forms is related to early spatial skills (Verdine, Bunger, Athanasopoulou, Golinkoff, & Hirsh-Pasek, 2017) and spatial skills are important for academic success, particularly in math and science (Newcombe, 2017). The causal link between shape knowledge and spatial skills is not likely conferred by just learning shape names, but rather occurs because shape exposure through play leads to children discussing the spatial properties of objects and exercising their spatial skills. For example, activities such as aligning shapes to be inserted into shape sorters or creating tangrams require spatial skills. These activities likely also foster attention to features of shape and lead to talk with parents about the number or sizes of sides on shapes and the similarities and differences between them. Because of the importance of shape knowledge and spatial skills in early math learning, it is important to consider ways of maximizing what children learn from their experiences with shapes.

1. Early shape learning

By age 3, children know the names of some basic shapes (Clements, Swaminathan, Hannibal, & Sarama, 1999; Verdine, Lucca, Golinkoff, Hirsh-Pasek, & Newcombe, 2016) but do not understand many of their defining properties, such as the number of sides or angles (Clements & Battista, 1992; Clements et al., 1999; Fisher, Hirsh-Pasek, Newcombe, & Golinkoff, 2013). Despite shape play being a common early childhood activity, it is not until

* Corresponding author at: University of Delaware, School of Education, Willard Hall, Newark, DE 19716, United States.
E-mail address: verdine@udel.edu (B.N. Verdine).

https://doi.org/10.1016/j.ecresq.2018.03.015
0885-2006/

elementary school that most children begin to accurately apply shape names to unusual variants, such as isosceles triangles, and learn that defective versions of shapes (e.g., shapes with gaps in the sides) are invalid category members (Satlow & Newcombe, 1998). This outcome is hardly surprising given the amounts and type of exposure they receive. For example, early educators in a university childcare center for children from birth to five years rarely used shape names or identified shape properties (Rudd, Lambert, Satterwhite, & Zaier, 2008). Young children are also rarely exposed to unusual variants that would challenge their concepts for shape categories. Shape toys seldom include varied versions from the same shape category and focus on canonical or "normal" shapes, such as equilateral triangles, circles, and squares (Resnick et al., 2016). Most shape sets do not include rectangles.

Yet there seems to be little reason to exclude complex or unusual shapes when teaching them and no sound developmental explanation for why it takes many children so long to learn the defining features of shapes. Preschool children are capable of learning shape categories, properties, and definitions to appropriately sort non-canonical shapes when taught using methods that highlight the defining features (Fisher et al., 2013). During a "guided play" interaction, an adult provided scaffolding to help children discover properties of shapes in a playful, exploratory manner. Children were told all of the shapes were "real" shapes and were asked to help in figuring out the shapes' secret (i.e., what makes them "real"). After exploring, the experimenter helped "discover" the distinguishing features through questions and tracing shapes as they examined exemplar cards.

This kind of teaching about shapes is likely not typical; even when teachers, who are presumably experts in educating children, talk about shapes, they focus on naming shapes rather than on their properties and characteristics (Sarama & Clements, 2004). Children quickly learn to attach the basic shape names to canonical versions; however, if there is no variety in the shapes being shown *and* little discussion of their properties, then repeated exposure and naming of canonical shapes simply reinforces incorrectly constrained ideas of what defines shape categories. Though we know of no data to support this suggestion, it is plausible that drilling children on canonical shapes names could cement overly narrow ideas about shapes and make it more difficult to later learn the actual defining features. Regardless, the existing literature strongly suggests that, rather than children being developmentally unprepared to learn more, it is probably how we are teaching shapes that leads many preschoolers to have immature or inaccurate shape knowledge.

2. Language and early spatial development

Parents often engage young children in play with shapes and shapes are relatively common in toy sets for infants and toddlers (Resnick et al., 2016). In Pruden, Levine, and Huttenlocher (2011), basic shape terms were among the common words parents used in 13.5 h of observation (circle was used by 80% of parents, shape by 72%, square by 65%, and triangle by 62%). However, mere exposure to shape names is likely insufficient for fostering deep shape knowledge *or* for spurring spatial skills and mastery of other spatial language. Terms describing the spatial features of shapes, for example, may be particularly crucial in bolstering children's shape knowledge and more likely to have effects on spatial skills.

Though most of the available studies are correlational in nature, hearing or using spatial language in informal contexts tends to be related to better spatial skills in children. For example, Pruden et al. (2011) observed parent spatial talk in the home and found that hearing spatial language during toddlerhood, including shape names, was related to spatial performance at 54 months on the Children's Mental Transformation Task (Levine, Huttenlocher, Taylor, & Langrock, 1999) and a Spatial Analogies test (adapted from Huttenlocher & Levine, 1990). Their mediation analyses suggest that, rather than a direct connection, these result from parent spatial talk encouraging child spatial talk, which is then related to spatial performance. Though there was no link between parent spatial talk and the WPPSI Block Design subtest after controlling for parents' non-spatial tokens, there was a direct relation of child spatial talk with the test.

The importance of spatial language for child spatial performance has also been documented in the context of shared book-reading. For example, parents of 5-year-olds who discussed spatial relationships between images that portrayed different viewpoints of the same scene had children with better spatial skills (Szechter & Liben, 2004). The importance of spatial language use has also been documented in a museum setting (Polinsky, Perez, Grehl, & McCrink, 2017), where an experimental manipulation of instruction to parents led to the use of more spatial language during a spatial construction task at an exhibit. Improvements in post-test scores on a spatial task were related to the amount of spatial language children used. In summary, children's use of spatial language during toddlerhood is related to a range of spatial skill improvements and their use of spatial language tends to be related to parents' use of spatial language. Therefore, finding ways to encourage both parent's and children's use of spatial language in early interactions is likely to improve later spatial skills for children.

Indeed there is general evidence that spatial activities are related to later spatial skill and mathematics performance (Levine, Ratliff, Huttenlocher, & Cannon, 2012; Lombardi, Casey, Thomson, Nguyen, & Dearing, 2018), including a number of studies focused on how differences in early spatial experience are likely involved (e.g., Nazareth, Herrera, & Pruden, 2013; Terlecki & Newcombe, 2005) in creating the male advantages in some spatial skills, notably mental rotation, which emerge after preschool (e.g., Frick, Möhring, & Newcombe, 2014; Linn & Petersen, 1985).

Early shape experience through joint parent–child play with shape toys is probably common already, given the availability of shape toys in the market. Parents and children are clearly talking about shapes with relative regularity (Pruden et al., 2011). Likewise, children appear to seek out play with shapes during free play. Seo and Ginsburg (2004) observed 4- and 5-year-olds from mixed economic backgrounds and found that the most common mathematical activities were categorized as being related to pattern and shape (21% of all instances; 8% higher than any other category). In play, adults and children can label shapes, talk about defining characteristics, and use spatial language or gestures to discuss shapes (Cartmill, Pruden, Levine, & Goldin-Meadow, 2010). But can we encourage dyads to do more, and to vary their input more?

3. The role of parents, play, and toy design

It is well established that the nature of parent–child interactions around different toys is predictive of related skills. For example, Ramani, Rowe, Eason, and Leech (2015) observed parent–child interactions during a semistructured play interaction using a series of three sets of materials chosen to elicit talk about numbers. They found that aspects of both parent's and children's math talk were related to the child's numerical knowledge. There is also evidence that seemingly simple changes to the design of learning materials like toys or games can influence what is learned from them. For example, Siegler and Ramani (2009) found that playing a linear board game increased understanding of numerical magnitudes more than a circular game.

There appears to be a role for spatial language in bolstering shape knowledge and shape play provides opportunities for increased exposure. A better understanding of how toy design influences parent–child interactions can guide decisions on how to improve play experiences without cumbersome interventions. We know that parents tend to use more spatial language with their children when engaged in spatially relevant play, but we also know that the amount they use varies widely (Ferrara, Hirsh-Pasek, Newcombe, Golinkoff, & Lam, 2011; Pruden et al., 2011). Investigating potential mechanisms for increasing the prevalence of spatial language in parent–child interactions is crucial. Even briefly providing parents with information on the significance of spatial language, and suggestions for how to incorporate spatial language or guidance into block play, can increase the amount of spatial language parents and children use during interactions (Borriello & Liben, 2017).

Since the introduction of the iPad in 2010, children are being offered many more virtual play experiences (Rideout, 2017). Though these virtual experiences afford some advantages over tangible materials, for example no physical limitations on the types of stimuli that can be included, research suggests that aspects of electronic instruction may be disruptive to task-relevant language compared to non-electronic toys. One study found that parents' language quality when playing with their child was greater when playing with a traditional shape sorter relative to an electronic one (Zosh et al., 2015). When playing with the electronic shape sorter, parents used fewer spatial terms and demonstrated fewer shape focused behaviors, instead talking about non-shape features of the toy. A similar pattern of results has been documented with 10- to 16-month-old infants during parent–child interactions with traditional toys, electronic toys, and books; adults produced fewer overall words and content-specific words with electronic toys than the other types (Sosa, 2016). These differences in the language produced by adults may be particularly important due to research showing that automated or electronic prompts may not be as effective for learning as those delivered by humans (e.g., Strouse, O'Doherty, & Troseth, 2013), especially for children with more limited vocabulary or executive function skills (Strouse & Ganea, 2016).

If electronic toys, which tend to closely resemble their non-electronic counterparts, can elicit substantially different reactions from parents and children, are there minor changes that can be made to toy designs that will substantially increase a focus on the defining features of shapes and exposure to spatial language other than just shape names? Research on concept formation (e.g., Carmichael & Hayes, 2001; Rakison & Oakes, 2003) and prior literature on shape learning (Satlow & Newcombe, 1998) suggests that designing shape toys with varied, less-canonical instances could provide better opportunities to learn their defining features. Toys with more varied shapes, including atypical variants, may increase children's understanding of shape properties and definitions (Resnick et al., 2016; Verdine et al., 2016). Furthermore, providing varied instances per category might also draw children's attention to their similarities and differences and provide opportunities to talk about these properties. Ultimately, careful design seems likely to increase spatial language talk about spatial and measurement terms such as relative length and size as well as providing practice with mathematics (e.g., counting sides). As Pruden et al. (2011) suggest, that type of spatial talk by children is specifically what links shape knowledge to spatial skills. The paucity of shape sets with these properties may be a proximal cause of children entering school without definition-focused concepts for most geometric shapes.

4. Research questions and hypotheses

The present study focuses on parent–child interactions with 3-year-old children during play with different geometric toys because, at this age, few children have deep knowledge of the defining features of geometric shapes. Therefore, optimal teaching of shapes for this age would likely be less about providing shape names and more about discussing the shapes' spatial properties and defining features. Touchscreen apps, which tend to repeat shape names, also tend not to reference shape properties. This may distract parents from talking about the spatial properties of shapes or simply reduce all shape conversation. Toys including a variety of shapes with different spatial properties that can be compared, on the other hand, may focus parents on those attributes and encourage dyads to discuss spatial properties and defining features.

Given these possibilities, our overarching question is whether the design of geometric toys influence parent and child language during play sessions focused on learning about shapes. Does shape naming, spatial language, or the number of questions parents and children ask naturally differ during play with shape toys that include more varied shapes (e.g., an equilateral *and* a scalene triangle) compared with toys that only contain standard, canonical versions? Does play with a shape-focused touchscreen app elicit similar parent–child interactions as does manual play with tangible shape sets containing the same shapes? If the design of shape materials impacts the extent to which adults and children discuss shapes or there are early gender differences in these interactions, it may have critical implications for educators, parents, and toy manufacturers.

We predicted that parents and children would use more spatial language and make more references to shapes with the alternate shape-toy set that included more varied shapes than a standard toy set. When presented with two shapes that look very different but are still considered, say, triangles, we expected these alternate shapes to provoke discussion of the similarities and differences between shapes and therefore increase the use of spatial and number-related language. We also expected the unusual shapes in the alternate condition might lead to confusion or curiosity from children or make it more likely that parents might try to lead their child to exploring the unusual properties of the shapes. Each of these possibilities we would expect to be accompanied by the groups asking more questions. Given prior work on decreased language quality and quantity in interaction with electronic versus traditional shape sorters (Zosh et al., 2015), we predicted that both parents and children would talk less overall and use less shape, spatial, and number-related language with a touchscreen tablet app than with a concrete set of the same shapes.

4.1. Exploratory analyses

Given general interest in gender differences for spatial skills and the fact that gender differences in some spatial skills emerge later, we did exploratory analyses including gender. If early shape experiences are important for spatial development and males tend to develop better spatial skills eventually, then we might expect to find that males and their parents use more or different spatial language during early interactions with shapes.

We also expected that the influence of including more varied shapes would result primarily in a general increase in spatial language use, rather than an increase specific in any particular type of spatial language. However, children or adults might focus very specifically on properties of the atypical shapes (e.g., the relative length of sides or other properties of the shapes), in which case, we might expect the spatial language increase to be largest for certain categories of spatial language. Therefore, we did include investigations of the spatial language subcategories as an exploratory set of analyses, where we would expect to see the biggest influence of the alternate shape condition on the number of words produced for the spatial dimensions (e.g., big, short, tall), continuous amount (e.g., less, part, more, piece), and spatial features and properties (e.g., angle, corner, line) subcategories.

5. Method

5.1. Participants

Participants were 60 parent–child dyads (child age $M = 35.8$ months; SD = 1.2; range = 34.3–38.1; females = 30; males = 30). Parents were overwhelmingly mothers ($N = 56$) but included fathers ($N = 3$) and grandparents ($N = 1$). All children were monolingual according to our criterion (>70% exposure to English – $M = 98.3$; $SD = 6.0$). Each of the three toy conditions contained 20 children (10 females). Participants were recruited by telephone through a database of volunteers to be contacted about participating in lab studies. The sample was predominantly white (52 Caucasian; 5 Black; 2 Other; 1 Unreported) and most parents had at least a 4-year degree (55 of 60). Children received a certificate, sticker, and t-shirt for their participation.

An additional 12 parent–child dyads came to the lab and were therefore randomly assigned to condition and consented to be in the study, but were ultimately dropped. 8 were dropped immediately following the sessions: 3 for the child having suspected developmental delays as judged from the questionnaire and low test performance; 2 for indicating less than 70% English exposure on the questionnaire; 2 for equipment problems that prevented data capture for coding; and 1 because the tester became not blind to condition during the procedure. An additional 4 were dropped as we made final coding decisions: 3 for not playing with the assigned materials for a minimum of 5 min; and 1 for a parent speech problem that made the session impossible to transcribe. There were 6 dropped participants in the tablet condition, 5 in the standard condition, and 1 in the atypical condition. The only participants dropped for a reason likely to have been influenced by assignment to condition were children dropped for not playing with the assigned material; given the common assumption that tablets are engaging, it is perhaps noteworthy that all 3 of the children dropped for not playing with the assigned materials had been randomly assigned to play with the tablet.

5.2. Materials

Each of the toys featured 10 shape categories for children to learn: circle, triangle, square, rectangle, hexagon, diamond, oval, crescent, star, and heart. These shapes and names were used because they appeared that way in the touchscreen app, which could not easily be altered. Though the names may not be the most mathematically appropriate (e.g., diamond instead of rhombus) or the shapes the most geometrically relevant (e.g., hearts and stars), the shapes sets are ecologically valid; these shapes and names are used often by shape toys (Resnick et al., 2016). With the exception of the occasional error (e.g., calling a hexagon an octagon), the only name that was commonly used by both parents and children that differed from that used by the tablet was calling the crescent a "moon."

The three possible toy sets, shown in Fig. 1, were: 20 wooden shapes including two canonical or "standard" versions of each of the 10 shape categories (the *standard* condition); 20 wooden shapes with one set of 10 standard shapes and one set of 10 alternate versions of each shape (the *alternate* condition); and an iPad game called Shapes Toddler Preschool by Toddler Teasers which randomly presented standard versions of the 10 shapes (the *tablet* condition).

The wooden shapes were made of 3/8" thick plywood painted yellow to match the tablet shapes. The standard sets of shapes were equilateral versions or had sides with "typical" proportions (e.g., the rectangle was close to a "golden rectangle" which have sides with a ratio of 1.62:1). They were modeled after the iPad app shapes to keep their appearances consistent with the tablet condition, but were scaled up to make them easier for children to handle. With the larger size, most shape sides increased approximately 1 cm in comparison to the tablet versions.

The alternate versions of the shapes varied the dimensions of the standard shapes to create versions that were different and yet still, by definition, valid members of their respective shape categories. In some cases, like with circle or square, the only way to alter the appearance without altering the shape category to which they belong was to alter the overall size. In the case of the rectangle, the scaling of orthogonal sides could be changed, but not the size of the angles. For many of the other shapes, the relative length of sides, angles, or the proportions of other parts could be altered. Controlling how "different" these shape sets were from one another was not possible because the shape alterations were different for each shape category.

The tablet used was a second-generation iPad (screen: 24.3 cm × 19.5 cm). The app (*Shapes Toddler Preschool*) featured 4 modes of play (quizzing, flashcards, toy box, and puzzle). Parents were asked to only play with the quizzing, flashcard, and puzzle modes, but were not directed how to allocate their time in those modes (see Fig. 1 for screenshots). The quizzing mode presented a word on the screen beneath a selection of shapes and said the shape name. When the child touched the correct shape the tablet would play a celebratory sound and proceed to the next shape. If the child touched the wrong shape, the tablet would indicate that it was wrong and the shape would disappear, leaving fewer options for the next attempt. The flashcard mode simply provided a page with the shape and name printed below it. The app would say the shape name when the shape appeared and children could "flick" through the shapes by swiping across the screen. The puzzle mode featured a shape at the bottom left and outlines of the shapes. The child had to drag the shapes into the outlines and, when successful, a new shape would appear until all the outlines were filled. Each time the child touched the shape to drag it, the tablet would say its name. A celebratory sound was played when the shape was properly placed. The toy box mode, which we asked parents not to use, featured many animated scenes in which music played. In piloting, we saw that this mode featured little shape naming content, parents and children tended to silently watch the action, there were few opportunities for parents to provide feedback, help, or interject additional information, and the mode could not be at all approximated with similar activities in the tangible material conditions.

5.3. Procedure

Dyads were randomly assigned to toy condition when scheduled to come to the lab. Parents were first given consent materials and a short questionnaire about their child's development. During this time children participated in a "warm-up" session in a playroom, which consisted of an undergraduate research assistant blind to the child's condition playing with the child using a set of general lab toys. All potentially geometric toys (e.g., magnetic building blocks) were specifically removed from the play area, leaving behind toys like a kitchen and cooking set, dolls, some figurines, books, and farm animals. The dyads then went to another quiet room where parents were instructed by an experimenter to play with the assigned toy for ten minutes while being video recorded (see full scripts in Appendices A and B). They were told their goal was to try to teach their child as many of the target shapes as possible using the assigned toy. In the tangible toy conditions, parents were offered suggestions for activities to complete that corresponded to the tablet app. In the tablet condition, parents were shown how to use

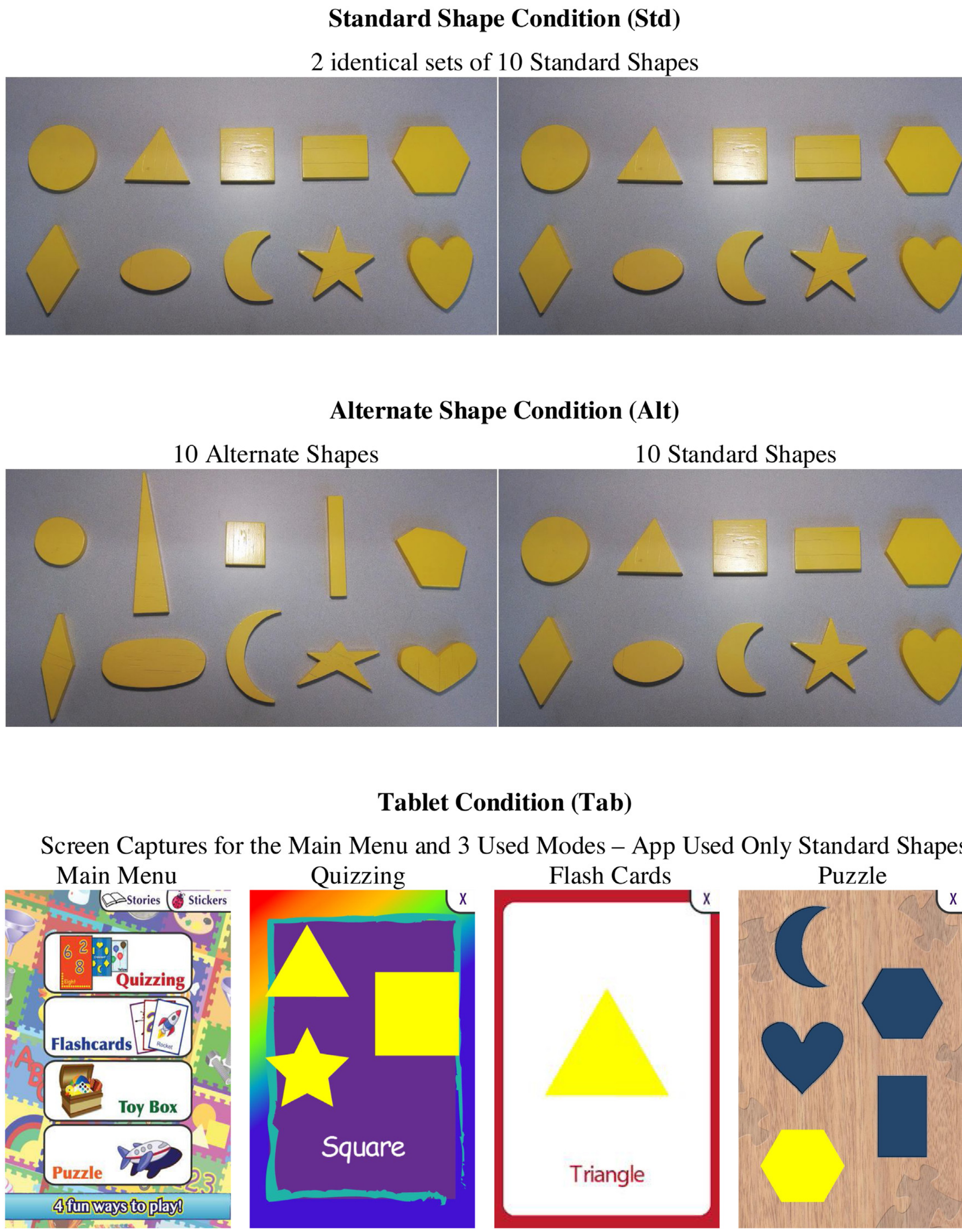

Fig. 1. The toys used for each of the 3 study conditions.

the tablet and the different modes of the app. An instructional guide on the wall was present for all conditions and provided an image and name for each shape. Parents were told to use this guide for their reference, but not directly use it to teach. The guide included the same shape names and images as the tablet app to try to ensure that all parents used the same shape names regardless of condition, since piloting had revealed that some dyads used alternate names (e.g., moon instead of crescent). The guide also ensured all parents knew the correct shape names.

5.4. Video processing

5.4.1. Transcribing

Parent, child, and tablet utterances from the first 5 min of each dyad's 10-min interaction were transcribed. This timeframe was selected to reduce the burden of transcribing videos and because much of the sample did not continue to play with the toys for the full 10 min without disruptions influencing the validity of the data (e.g., bathroom breaks, seriously off-topic conversations). Concerns about data validity would have also been differential across conditions; more children in the tablet condition tended to be distracted during the final 5 min. Indeed the only 3 children dropped for not playing in the first 5 min were in the tablet condition. This decision to focus on the first 5 min, if anything, likely reduced the effect sizes we would have otherwise observed for comparisons to that condition.

Transcriptions were coded and evaluated using CLAN software (MacWhinney, 2000). All reported reliability statistics are Krippendorff's alpha, a statistic described in Hayes and Krippendorff (2007) with macros for calculating it. The advantage to using Kripendorff's alpha is that it allows reliabilities to be calculated and compared across any numbers of coders, values, different metrics, and with unequal sample sizes, providing a means to calculate a single reliability statistic for all of the coded variables that can be more readily compared within and across studies. Alpha levels above 0.800 are considered acceptable with those from 0.667 to 0.800 recommended for more tentative conclusions (Krippendorff, 2004).

Because of the amount of video to transcribe, 11 transcribers were initially used. A manual was created by the first author, with input from the lab manager and senior undergraduates. Once written, the manual was used to practice transcribing a pilot participant, which the lab manager checked for general accuracy and total word count. After meeting to discuss concerns, the transcriber continued to a second practice transcription. After a second meeting, transcribers were randomly assigned to participants and each interaction was transcribed twice. This approach allowed reliability assessment of each transcriber against a random set of other transcribers and ensured we could identify transcribers with poor reliability or difficult videos to transcribe. The only edits made to transcriptions, once they were completed and prior to reliability analysis, were for spelling corrections and typos. The changes were made without comparison between transcription pairs by the first author using word replacements in CLAN.

Three transcribers completed transcriptions but had low reliability. Their allotment of participants was re-transcribed by those with higher reliability and only data from the 8 transcribers with acceptable reliability are reported. The 3 dropped transcribers consistently undercounted tablet word counts; the tablet would repeat a word each time the child touched a shape, which they sometimes did rapidly, making them difficult to accurately count. Once they were removed, reliability was good and all remaining discrepancies for pairs of transcriptions were small, so transcriptions were not further rectified.

The average reliability between the retained transcribers was high for the total number of words (*M* Krippendorf's $\alpha = 0.98$; range = 0.96–0.99), utterances (*M* Krippendorf's $\alpha = 0.97$; range = 0.96–0.99), tagged spatial words (*M* Krippendorf's $\alpha = 0.97$; range = 0.96–0.99), and questions (*M* Krippendorf's $\alpha = 0.79$; range = 0.68–0.89). Because of acceptable reliability, from each transcription pair a "master" transcription was selected, which was always the one transcribed by the coder with the higher average reliability. This master transcription was then used for all further language counts and for utterance-level coding.

5.4.2. Language counts

CLAN software generated counts for all *tokens* and *types* for each speaker in a transcription. CLAN was also used to count the total number of shape names, spatial words, number words, and questions. *Tokens* to refer to the total number of words spoken, regardless of whether a word had been used previously in the interaction. *Types* are the number of unique words spoken. For example, if a speaker said *triangle* three times, triangle would be counted three times as a shape name token and three times for all tokens, but only counted once for all types and shape word types. Groups of specific word types were counted by generating a list to be searched out (see Appendix C), which CLAN then counted for each speaker in every master transcription. In general, our strategy for choosing words was to be more exhaustive in counting than may seem entirely necessary for each category. If the additional words were truly unnecessary and did not actually occur, then they would not be counted and have no influence on the statistics.

5.4.2.1. Shape names. The list of shape names was first generated from the shapes section of the Cannon, Levine, and Huttenlocher (2007) spatial language coding manual. Our lab generated additional names to make a more exhaustive list (see Appendix C). Though very rarely used, the additional names were included to capture any potential talk about more advanced shapes.

5.4.2.2. Spatial words. The list of spatial words was generated by using each spatial word listed in 7 of the 8 categories defined in Cannon et al. (2007). This list excluded the category of shape names because they were counted separately and the main spatial language variable included only the other 7 spatial subcategories. For exploratory purposes, the 7 spatial word subcategories were tracked separately and analyzed, though the word counts for some categories were quite low and analyses related to these subcategories should be interpreted with some caution.

5.4.2.3. Number words. Number words were counted for the numbers 1–20 plus 50 and 100. This list includes the numbers necessary for describing features of the shapes (e.g., 4 sides) and all of the numbers that dyads would need to count all of the shapes in the shape sets. Fifty and 100 were never actually used and reference to numbers higher than 10 were quite rare, which is not surprising since most early 3-year-olds are not yet adept at counting and few truly understand the cardinal principle (Sarnecka & Carey, 2008).

5.4.2.4. Questions. Transcribers marked utterances that were questions while they were transcribing and these were counted from the master transcriptions. Coders were allowed to use intonation and context to infer whether a statement could be considered a question and most ambiguity centered on grammatically incomplete child statements.

6. Results

6.1. Statistical approach

Child production, parent production, and all language exposure (parent + tablet) are the 3 "speaker categories" used throughout the analyses. The all language exposure variables count language spoken by the parent and the tablet together (i.e., all heard language); these analyses use only the parent language counts for the alternate and standard conditions, since those conditions had no other sources of language heard by the child. Within this paper, we focus our analyses and discussion on token counts as opposed to type counts. This is typical of prior research on interactions (e.g., Pruden et al., 2011) and, in our data, patterns were generally similar for tokens and types. Analyses including types are available in the online supplement. For most tables in the results, abbreviations are used for condition (i.e., the toy type that was being played with): Tab = tablet condition; Std = standard shape condition; and Alt = alternate shape condition.

Since the dependent variables are counts, many distributions were non-normal and homogeneity of variance was a problem. Independence could also not be assumed since most dependent variables were subsets of the all language variables. Therefore, many parametric statistics were not appropriate (e.g., MANOVA) and instead bootstrapping was used. Bootstrapped parameters were estimated using SPSS with 10,000 iterations resampled with replacement from the original dataset. Resampling was done for all means and each individual regression.

The main analyses in the first section of the results were guided by prior research and targeted to answer our specific research questions. These regressions included dummy variables for the alternate condition and tablet condition as predictors of the 5 dependent variables (all tokens, shape name tokens, spatial tokens, number tokens, and questions). 4 of these regressions were repeated using the 5th dependent variable, all tokens, as a covariate. The regressions with the all language covariate entered prior to the predictors tell us whether each predictor variable significantly predicts the number of those specific words or questions, taking into account the fact that speakers who talk more overall are likely to produce more questions and more of every word type.

For the purposes of generating a complete report and informing future research, we follow this section with a section presenting analyses that include other variables (e.g., gender) and that investigated the subsets of our spatial language variable. We consider these analyses and results exploratory, and therefore in need of replication and targeted study. They appear in their entirety in the online supplement.

All regressions use the standard condition as the reference sample, with β weights relative to that condition; positive values indicate higher production relative to the standard shape set and negative values indicate lower. Because the variables are all counts and the condition and sex variables are dummy coded (1 and 0), the β weights "units" for condition and sex can be interpreted as words or questions. For example, if the Tablet condition had a β weight of −12.0 for child shape name tokens, a child assigned to play with the tablet would be expected to produce 12 fewer shape name tokens during a 5-min interaction than a child assigned to play with the standard shape set. The standard condition was the reference sample for theoretical and practical reasons; the best comparison toy for the alternate and tablet materials was the standard shape set, sharing the same shape designs with the tablet and sharing the same physical form with the alternate shapes. We also predicted that the standard condition means would be intermediate between the other conditions, a pattern the data tend to follow, yielding more easily interpreted β weights which are mostly positive for the alternate condition and mostly negative for the tablet. Use the 95% confidence intervals for the alternate and tablet condition β weights to compare those groups.

6.2. Main study analyses

The mean number of tokens produced by children, parents, and the parents and tablet together are presented in Fig. 2 and Table 1. Full descriptive statistics for the main study variables including raw means, percentiles (Table S1), and bootstrapped means (Table S2) are reported in the online supplement. The overall regression model statistics are reported in Table 2 for each of the 3 speaker groups. The *p*-value cutoff for a Bonferroni correction of 5 repeated regression analyses is $p < 0.0102$. Therefore, any analysis in Table 2 with two asterisks meets that cutoff. For ease of interpretation, especially for the all language exposure analyses, regardless of significance, we have chosen to report individual regression estimates (Tables 3 and 4) for the main dependent variables without a covariate (models labeled C). The ΔF *p*-values from Table 2 appear in brackets in the model columns of Tables 3–5. The models using all tokens as a covariate (Tokens + C) were only included when the overall model met the $p < 0.01$ criteria.

In most of the regressions with just the condition variables entered (labeled C in Table 2), parent and child production were influenced by condition assignment, with the exception of number tokens for the children. Outcomes for the regressions using data which included language produced by the tablet (all language exposure) are quite different from the child and parent production regressions and only shape names and numbers have significant condition differences. Not surprisingly, the all tokens covariate tended to be a significant predictor in regressions where the covariate was entered; the only exception was for the regression predicting shape name tokens in the Parent + Tablet condition (Table 2). This exception is likely because the tablet produced an exceptionally large number of shape name tokens (about 65 on average; Fig. 2) and therefore an unusually large percentage of shape names as a function of overall token production. These patterns in the overall regression models are explored as we consider the condition and covariate regression estimates for child and parent production (Table 3) and for parent + tablet production (Table 4).

6.2.1. Child production

As predicted, children in the alternate shape set condition produced significantly more overall tokens ($\beta = 43.7$; Std. $\beta = 0.39$; $p = 0.001$), shape names ($\beta = 4.9$; Std. $\beta = 0.23$; $p = 0.047$), and spatial language tokens ($\beta = 7.7$; Std. $\beta = 0.55$; $p = 0.001$) than the standard shape set and the tablet. They did not produce a significantly different amount of number tokens ($\beta = -1.2$; Std. $\beta = -0.06$; $p = 0.720$) or questions ($\beta = 4.0$; Std. $\beta = 0.32$; $p = 0.102$). The regression predicting spatial language tokens using all tokens as a covariate was marginally significant for children but was excluded to correct for multiple comparisons ($p = 0.023$). However, as expected in that regression, children in the alternate condition did appear to produce more spatial language tokens after accounting for all tokens ($\beta = 3.9$; Std. $\beta = 0.28$; $p = 0.016$). Therefore, there is strong evidence that children playing with the alternate shapes produced more spatial words overall compared to those playing with only standard shapes and there is somewhat weaker evidence that they produced more spatial words even after accounting for their overall language use.

Also as predicted, children in the tablet condition tended to produce fewer words than children playing with the tangible shape sets; they produced fewer overall tokens ($\beta = -51.3$; Std. $\beta = -0.46$; $p < 0.001$), shape names ($\beta = -9.1$; Std. $\beta = -0.43$; $p = 0.002$), and questions ($\beta = -2.2$; Std. $\beta = -0.18$; $p = 0.034$).

6.2.2. Parent production

Production of the parents in the standard and alternate shape conditions was similar for all of the main study variables. However, for all 5 of the main variables, parents did produce fewer instances in the tablet condition compared to the standard condition [all tokens ($\beta = -131.2$; Std. $\beta = -0.45$; $p < 0.003$), shape name tokens ($\beta = -12.0$; Std. $\beta = -0.33$; $p = 0.021$), spatial language tokens ($\beta = -16.8$; Std. $\beta = -0.40$; $p = 0.001$), number tokens ($\beta = -9.7$; Std. $\beta = -0.47$; $p = 0.001$), and questions ($\beta = -16.4$; Std. $\beta = -0.55$; $p = 0.001$)]. The 95% confidence intervals for the alternate and tablet condition, which do not overlap with the β weights for each other, indicate that these conditions were also significantly different from each other at a $p < 0.05$ level.

6.2.3. All language exposure analyses

Comparing the two tablet count means in Fig. 2 or Table 1 (with and without tokens from the tablet), we can see that on average the tablet added an appreciable number of overall tokens (132.9), specifically adding many shape name tokens (65.5), with some additional spatial language (6.3) and questions (6.5). Note that the tablet did repeatedly produce a single deictic spatial word (where) that was used in the context of a single question prompt for one of the modes ("Where is the [shape name]?"). This prompt accounts entirely for the 6.3 spatial language tokens contributed by the tablet, accounts for nearly all of the phrases that were coded as

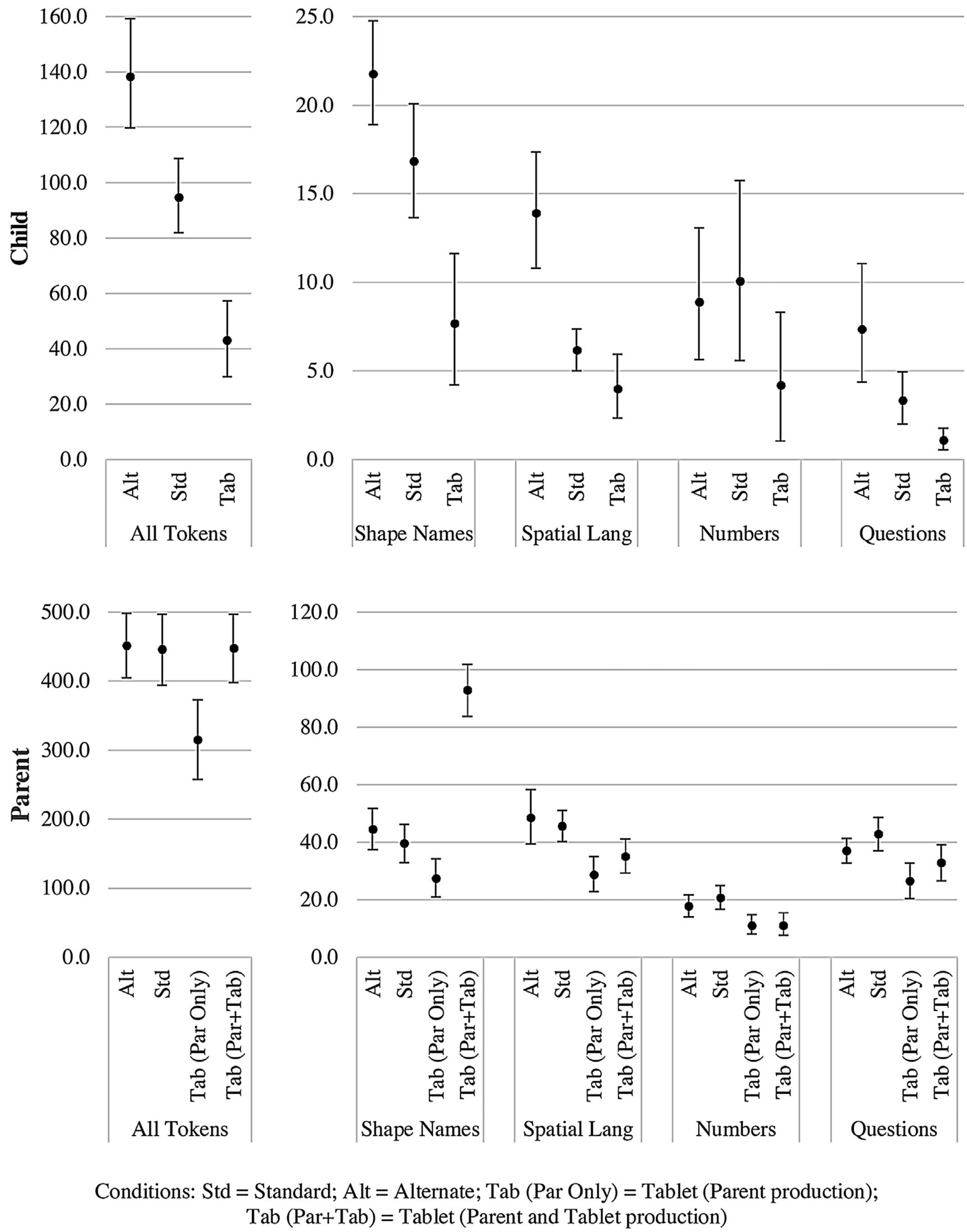

Fig. 2. Plots showing bootstrapped mean tokens with 95% confidence intervals by speaker for the main study variables (all tokens, shape name tokens, spatial language tokens, number tokens, and questions asked).

questions, and was the source of about 1/10th of the shape names contributed by the tablet.

Considering the production from the tablet, there should be little surprise that the all language exposure regressions in Table 4 show no condition differences for all tokens; the number of overall tokens the tablet produced almost exactly matched how many less words the parents produced in the tablet condition relative to the others. Because the vast majority of these words were shape name tokens, the tablet condition heard many more shape names ($\beta = 53.5$; Std. $\beta = 0.85$; $p < 0.001$; all tokens covariate; $\beta = 53.5$; Std. $\beta = 0.85$; $p < 0.001$) than both of the other two conditions. The small amount of spatial language produced by the tablet was enough to make the spatial language regression not significant ($p = 0.051$), even though it was in the parent production analyses. However, because the tablet added a lot of words that were not spatial, the tablet condition heard fewer spatial words after accounting for all tokens ($\beta = -9.7$; Std. $\beta = -0.48$; $p < 0.001$). The tablet did not produce number words so there was no change relative to the parent production analyses for the regression including only the conditions as predictors ($\beta = -9.6$; Std. $\beta = -0.47$; $p = 0.001$; minor variances compared to the parent production table are due to the bootstrapping). However, again because other types of words were added, the regression controlling for all tokens did come out significant; number words heard in the tablet condition were lower

Table 1
Bootstrapped Means by Condition and Sex with 95% Confidence Intervals for Main Study Variable Tokens.

		Alt			Std			Tab			Female			Male		
		M	95% CI		*M*	95% CI		*M*	95% CI		*M*	95% CI		*M*	95% CI	
			Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi		Lo	Hi
Child tokens	All tokens	138.3	119.7	159.2	94.6	81.8	108.7	43.2	29.9	57.4	92.4	74.2	110.3	91.8	71.7	113.7
	Shape names	21.8	18.9	24.8	16.8	13.7	20.1	7.7	4.2	11.6	16.8	12.7	20.8	14.0	10.9	17.2
	Spatial lang	13.9	10.8	17.4	6.2	5.0	7.4	4.0	2.4	6.0	8.9	6.4	11.8	7.1	5.2	9.4
	Numbers	8.9	5.7	13.1	10.1	5.6	15.8	4.2	1.1	8.3	7.1	4.8	9.7	8.3	4.9	12.4
Parent tokens	All tokens	451.3	404.6	498.6	446.0	394.2	497.2	314.6	257.4	373.1	381.8	339.9	421.7	426.5	369.5	480.7
	Shape names	44.4	37.5	51.9	39.5	32.9	46.2	27.4	20.8	34.1	34.7	28.7	41.2	39.5	33.2	45.4
	Spatial lang	48.5	39.4	58.3	45.6	40.3	51.1	28.7	22.9	35.0	35.1	30.2	40.3	46.8	39.4	54.3
	Numbers	17.7	14.0	21.6	20.6	16.7	24.9	11.0	8.0	14.7	14.6	11.8	17.7	18.3	14.9	21.7
Parent + tab tokens	All tokens	–	–	–	–	–	–	447.5	397.6	497.3	427.5	392.8	461.5	469.1	420.7	517.3
	Shape names	–	–	–	–	–	–	92.9	83.7	101.9	56.2	45.4	67.7	61.6	51.4	72.1
	Spatial lang	–	–	–	–	–	–	35.0	29.2	41.0	37.5	32.4	43.0	48.6	41.0	56.7
	Numbers	–	–	–	–	–	–	11.0	7.6	15.4	14.6	11.6	17.8	18.3	14.7	22.0
Questions	Child	7.3	4.4	11.1	3.3	2.0	5.0	1.1	0.6	1.8	4.6	2.5	7.2	3.3	2.2	4.5
	Parent	37.0	32.7	41.3	42.7	37.1	48.6	26.3	20.3	32.7	34.2	29.4	39.0	36.5	31.7	41.3
	Parent + tablet	–	–	–	–	–	–	32.8	26.5	39.2	36.6	32.3	40.9	38.4	33.7	42.9

Table 2
Overall regression model statistics.

Dep variable	Model[a]	*df*1	*df*2	Child production			Parent production		All language exposure[b]		
				Adj. R^2	ΔF	$\Delta F\ p$	ΔF	$\Delta F\ p$	Adj. R^2	ΔF	$\Delta F\ p$
1. All tokens	C	2	57	0.52	33.01	0.000[**]	7.82	0.001[**]	−0.03	0.01	0.988
2. Shape names	C	2	57	0.32	14.64	0.000[**]	5.73	0.005[**]	0.65	55.89	0.000[**]
	Tokens	1	58	0.47	52.39	0.000[**]	48.58	0.000[**]	0.04	3.29	0.075
	Tokens + C	2	56	0.46	0.92	0.405	1.00	0.374	0.70	65.76	0.000[**]
3. Spatial language	C	2	57	0.38	19.20	0.000[**]	6.83	0.002[**]	0.07	3.14	0.051
	Tokens	1	58	0.56	75.86	0.000[**]	81.72	0.000[**]	0.44	47.40	0.000[**]
	Tokens + C	2	56	0.60	4.02	0.023[*]	0.75	0.479	0.52	5.88	0.005[**]
4. Numbers	C	2	57	0.03	1.99	0.145	6.12	0.004[**]	0.15	6.12	0.004[**]
	Tokens	1	58	0.11	8.56	0.005[**]	30.55	0.000[**]	0.19	14.42	0.000[**]
	Tokens + C	2	56	0.12	1.13	0.331	1.76	0.182	0.34	7.95	0.001[**]
5. Questions	C	2	57	0.16	6.70	0.002[**]	8.78	0.000[**]	0.07	3.15	0.050
	Tokens	1	58	0.16	11.87	0.001[**]	98.79	0.000[**]	0.49	58.13	0.000[**]
	Tokens + C	2	56	0.17	1.48	0.235	3.57	0.035[*]	0.58	7.27	0.002[**]

[*] $p < 0.05$.
[**] $p < 0.01$ (Bonferroni-corrected significant level).
[a] Models marked C used two condition dummy variables as predictors to predict effects of condition. Models marked "Tokens" entered the all tokens variable as the only predictor. Models marked "Tokens + C" are stepwise regressions which entered all tokens as the first predictor (covariate) and then added the condition variables. For the all tokens dependent variable (1), regressions could not be run entering all tokens as a predictor.
[b] In the all language exposure analyses, the tablet condition includes production from both the parent and tablet. The standard and alternate shape sets only include production from the parent, since no tablet was present and there were no other sources of language for the child to hear.

relative to the other conditions after accounting for all tokens ($\beta = -9.7$; Std. $\beta = -0.48$; $p < 0.001$). Similar to the spatial language tokens regression, the small number of questions produced by the tablet was enough to cause a change compared to the parent production analyses and the spatial language regression was marginal ($p = 0.050$). Again though, children in the tablet condition heard fewer questions after controlling for overall tokens ($\beta = -9.9$; Std. $\beta = -0.37$; $p < 0.001$).

To summarize, children in the tablet condition heard about the same amount of overall language as those in the other conditions, but a significant amount of that language was produced by the tablet and a very large amount of the language it contributed was restricted to 10 shape names. This means that those in the tablet condition actually heard more shape names than the other conditions but still tended to hear less spatial language, numbers, and questions, after controlling for all tokens. Likewise, the little spatial language and few questions the tablet did contribute were due to a single repeated prompt ("Where is the [shape name]?").

6.3. Exploratory analyses

The exploratory regression analyses are reported in full in the online supplement. They used the two condition dummy variables but also included, sex (males = 1), and two condition by sex interaction terms (sex × alternate and sex × tablet) as their standard set of predictors. These analyses also included regressions predicting each of the spatial language subcategories.

6.3.1. Sex differences

For the child production regressions including sex, none of the β weights for sex and none of the sex by condition interactions were significant, meaning that male and female children produced similar amounts of language of all types regardless of condition, as expected. However, overall effects of the sex of the child on parent word production are evident for shape name tokens, with boys hearing more across condition ($\beta = 12.9$; Std. $\beta = 0.37$; $p = 0.048$). Though this is marginal, there are also significant Sex × Alternate condition interactions for the regressions predicting all spatial language types ($\beta = 6.3$; Std. $\beta = 0.41$; $p = 0.038$; all language covaried – $\beta = 4.8$; Std. $\beta = 0.32$; $p = 0.039$) and tokens and types for the continuous amount spatial subcategory (tokens $\beta = 8.7$; Std. $\beta = 0.71$; $p = 0.017$; types $\beta = 3.2$; Std. $\beta = 0.68$; $p = 0.003$; all language covaried – tokens $\beta = 9.0$; Std. $\beta = 0.73$; $p = 0.009$; types $\beta = 2.8$; Std. $\beta = 0.60$; $p = 0.002$). The graphs in Fig. 3 and means in Table 1 help illustrate the pattern that emerges from these data; there is a weak but general trend toward parents producing more language

Table 3
Condition regression estimates for main study variables for child and parent production.

Dep Variable	Model[b]	Ind variable	Bootstrapped parameters[a]					Std. β	t	Zero-order r
			β	SE	p	CI Low	CI high			
Child production										
1. All tokens	C [<0.001]	Constant	94.5	7.33	0.000	81.32	108.48		11.44	
		Alt	43.7	12.17	0.001	19.74	68.46	0.39	3.73	0.62
		Tab	−51.3	10.35	0.000	−72.01	−30.55	−0.46	−4.39	−0.65
2. Shape names	C [<0.001]	Constant	16.8	1.86	0.000	13.32	20.50		9.02	
		Alt	4.9	2.38	0.047	0.14	9.60	0.23	1.88	0.45
		Tab	−9.1	2.85	0.002	−14.51	−3.46	−0.43	−3.45	−0.55
3. Spatial language	C [<0.001]	Constant	6.1	0.61	0.000	5.00	7.32		5.17	
		Alt	7.7	1.76	0.001	4.51	11.22	0.55	4.61	0.62
		Tab	−2.1	1.16	0.077	−4.30	0.25	−0.15	−1.28	−0.42
4. Numbers	C [.145]	Constant	10.1	2.65	0.002	5.73	15.44		4.59	
		Alt	−1.2	3.31	0.720	−8.46	5.53	−0.06	−0.39	0.08
		Tab	−5.8	3.25	0.082	−12.60	0.38	−0.28	−1.89	−0.25
5. Questions	C [.001]	Constant	3.3	0.75	0.001	2.10	4.80		2.74	
		Alt	4.0	2.10	0.102	0.51	8.30	0.32	2.31	0.41
		Tab	−2.2	0.83	0.034	−4.09	−0.59	−0.18	−1.30	−0.34
Parent production										
1. All tokens	C [.001]	Constant	446.0	26.45	0.000	394.40	498.17		16.14	
		Alt	4.9	36.13	0.881	−64.49	73.85	0.02	0.14	0.25
		Tab	−131.2	40.56	0.003	−211.44	−51.27	−0.45	−3.35	−0.46
2. Shape names	C [.005]	Constant	39.5	3.37	0.000	32.84	46.00		10.83	
		Alt	4.9	5.07	0.345	−5.17	15.25	0.13	0.94	0.29
		Tab	−12.0	4.97	0.021	−21.68	−2.38	−0.33	−2.35	−0.39
3. Spatial language	C [.002]	Constant	45.5	3.07	0.000	39.56	51.47		11.12	
		Alt	3.0	6.16	0.627	−8.44	14.99	0.07	0.52	0.27
		Tab	−16.8	4.56	0.001	−25.21	−8.17	−0.40	−2.91	−0.44
4. Numbers	C [.004]	Constant	20.7	2.07	0.000	16.81	24.77		10.32	
		Alt	−2.9	2.87	0.323	−8.74	2.82	−0.14	−1.02	0.09
		Tab	−9.7	2.80	0.001	−14.99	−4.05	−0.47	−3.41	−0.40
5. Questions	C [<0.001]	Constant	42.8	2.86	0.000	37.05	48.73		15.25	
		Alt	−5.8	3.66	0.130	−12.86	1.18	−0.19	−1.44	0.08
		Tab	−16.4	4.24	0.001	−24.40	−8.26	−0.55	−4.13	−0.46

Note. The standard shape condition is the reference condition with Beta weights for the other conditions relative to the standard shape set (Alt = alternate; Tab = tablet).
[a] Based on 10,000 samples; confidence intervals (CIs) are bias corrected and accelerated (BCa).
[b] C models used dummy variables for the alternate condition and tablet condition as predictors. Models marked "Tokens + C" entered all tokens as a covariate prior to entering the condition predictors. Values appearing in brackets below the models are the ΔF p-values from the overall regression statistics.

with boys when playing with shape toys, but the alternate shapes, specifically, appear to have some impacts on the language parents produce that is dependent on the sex of their child.

6.3.2. *Spatial language subsets*

Spatial subcategory D, orientation and transformation words, had means that were effectively 0 for all speakers and all conditions. Subcategory H, pattern words, were also almost 0 for all child production conditions and were very low for all parent production conditions ($M < 1$). Therefore, these variables had a lack of variability and, not surprisingly, none of the analyses using them were significant.

As previously reported, children in the alternate condition produced more overall spatial language. The only subcategory regression for tokens that was significant was the regression predicting category F- Deictic tokens ($\beta = 5.3$; Std. $\beta = 0.77$; $p = 0.029$), which includes the words: anywhere, everywhere, here, nowhere, somewhere, there, where, and wherever.

Parents produced fewer spatial language tokens in the tablet condition. The specific subcategories for which parents in the tablet condition produced fewer spatial tokens relative to the standard shape set were category C- Location and Direction ($\beta = -10.9$; Std. $\beta = -0.45$; $p = 0.001$) and category E- Continuous Amount ($\beta = -3.3$; Std. $\beta = -0.34$; $p = 0.011$).

Analyses for all language exposure analyses (Parent + Tablet) were similar to the parent production analyses, except that the regressions with the covariate for category C and E also came out significant; the tablet contributed many overall words, but none in those specific categories, so when tablet language is included in the analysis the number of words from those categories is lower after controlling for all language. Because of the repeated prompt from the tablet noted previously ("Where is the [shape name]?"), the tablet did produce a single deictic spatial word (where) that happened with enough frequency that children in the tablet condition heard more deictics than the standard condition when tablet language was included ($\beta = 8.6$; Std. $\beta = 0.64$; $p = 0.001$; with all types covariate $\beta = 8.2$; Std. $\beta = 0.61$; $p = 0.002$).

7. Discussion

The design of geometric toys influences parent and child language during interactions around those toys. Seemingly subtle differences had critical effects on the amount of language both parents and children used. Several of our initial hypotheses were supported. Children use more overall tokens, more shape names, and more spatial language tokens with shape toy sets when they include alternate versions of the shapes. Because it is the child's language that appears to be the proximate predictor for later spatial skills (Pruden et al., 2011), these are important effects. They suggest interactions can be enriched with more spatial language and more meaningful conversations about shapes by making only slight adjustments to toys that many toddlers already use (e.g., shape sorters). Furthermore, these desirable outcomes were achieved without prompting or other intervention. The affordances of the shape toys included in the set are clearly spurring more spatial language that contains information about the key properties of shapes.

Parents, on the other hand, did not as strongly alter their language production with a toy set including more varied shapes as

Table 4
Condition regression estimates for main study variables for all language exposure (tablet condition includes parent production and tablet production).

Dep variable	Model[b]	Ind variable	Bootstrapped parameters[a]					Std. β	t	Zero-order r
			β	SE	p	CI Low	CI high			
1. All tokens	C [.998]	Constant	445.7	26.44	0.000	394.50	498.15		17.09	
		Alt	6.0	35.82	0.882	−65.40	77.34	0.02	0.15	0.02
		Tab	1.9	37.37	0.967	−69.96	75.30	0.01	0.04	−0.01
2. Shape names	C [<0.001]	Constant	39.5	3.34	0.000	33.20	46.11		9.99	
		Alt	4.9	5.04	0.337	−4.87	14.97	0.08	0.87	−0.35
		Tab	53.5	5.59	0.000	42.28	64.53	0.85	9.56	0.81
	Tokens + C [<0.001]	Constant	12.1	8.11	0.135	−4.00	28.50		1.35	
		All tokens	0.1	0.02	0.001	0.03	0.10	0.23	3.30	0.23
		Alt	4.6	4.43	0.318	−3.82	13.21	0.07	0.87	−0.35
		Tab	53.5	5.19	0.000	43.01	63.44	0.85	10.34	0.81
3. Spatial language	C [.051]	Constant	45.5	3.10	0.000	39.32	51.79		11.30	
		Alt	3.0	6.13	0.631	−8.38	14.54	0.08	0.53	0.21
		Tab	−10.6	4.42	0.023	−18.82	−2.33	−0.27	−1.86	−0.31
	Tokens + C [.005]	Constant	−3.0	7.39	0.693	−18.74	11.03		−0.41	
		All tokens	0.1	0.02	0.000	0.08	0.14	0.67	7.41	0.67
		Alt	2.2	4.51	0.600	−6.18	10.78	0.06	0.59	0.21
		Tab	−10.8	3.37	0.003	−17.64	−4.15	−0.27	−2.63	−0.31
4. Numbers	C [.004]	Constant	20.6	2.05	0.000	16.76	24.88		10.32	
		Alt	−2.9	2.84	0.320	−8.57	2.71	−0.14	−1.02	0.09
		Tab	−9.6	2.80	0.001	−14.83	−4.14	−0.47	−3.41	−0.40
	Tokens + C [.001]	Constant	3.8	3.94	0.338	−4.19	11.24		0.88	
		All tokens	0.0	0.01	0.000	0.02	0.06	0.45	4.23	0.45
		Alt	−3.1	2.49	0.225	−7.99	1.88	−0.15	−1.25	0.09
		Tab	−9.7	2.39	0.000	−14.52	−4.80	−0.48	−3.90	−0.40
5. Questions	C [0.050]	Constant	42.7	2.86	0.000	37.38	48.33		15.26	
		Alt	−5.6	3.67	0.133	−13.07	1.45	−0.21	−1.44	−0.03
		Tab	−9.9	4.18	0.024	−18.23	−1.88	−0.36	−2.50	−0.26
	Tokens + C [0.002]	Constant	6.9	4.58	0.134	−2.82	15.43		1.49	
		All tokens	0.1	0.01	0.000	0.06	0.10	0.71	8.44	0.71
		Alt	−6.0	2.70	0.028	−11.33	−0.78	−0.23	−2.32	−0.03
		Tab	−9.9	2.65	0.000	−15.29	−4.60	−0.37	−3.78	−0.26

Note. The standard shape condition is the reference condition with Beta weights for the other conditions relative to the standard shape set (Alt = alternate; Tab = tablet).
[a] Based on 10,000 samples; confidence intervals (CIs) are bias corrected and accelerated (BCa).
[b] C models used dummy variables for the alternate condition and tablet condition as predictors. Models marked "Tokens + C" entered all tokens as a covariate prior to entering the condition predictors. Values appearing in brackets below the models are the ΔF *p*-values from the overall regression statistics.

we had expected. Despite affording parents, the opportunities to highlight defining properties of the shapes, they do not appear to have done so in many cases. Anecdotally, at least one of the reasons may be that many parents seemed a little unsure of how to conceptualize some of the alternate shapes themselves. There were also a somewhat surprising number of parents who used the wrong shape terms entirely (e.g., octagon for hexagon). Other studies have shown that simple sets of instructions to parents can influence their subsequent language production and increase their use of spatial language, which we might expect to "trickle down" to children using even more of this language. Though the purpose of this study was to see what parents and children would do without specific instruction, a brief intervention or written explanation of the reasoning behind including the unusual shapes would likely increase parents' attention to defining features and the number of shape comparisons they make. Previous research has demonstrated that spatial language exposure in early childhood may be an important factor for spatial performance (Pruden et al., 2011; Szechter & Liben, 2004) and future research can continue to try to find ways to optimize that exposure.

The importance of language and the impact of early spatial experiences on skills are well established (e.g., Fisher et al., 2013; Pruden et al., 2011). Though this study does not link the observed language differences to actual child outcomes, papers have documented the role of parental language in teaching spatial information (e.g., Levine et al., 2012; Pruden et al., 2011). As Roger Brown noted long ago (1958), using language invites the formation of categories. Spatial categories might be broadened and refined when parents and children are presented with stimuli that include non-standard shapes that appear to invite comparison and discussion. Furthermore, when two things receive the same label (e.g., a long skinny rectangle and a shorter, fat one both being called rectangle), children may compare them, noting which features are common and which are not (Gentner & Namy, 1999; Markman, 1989; Waxman & Hall, 1993). As Gentner (2016) has argued, having the opportunity to compare two similar things invites discovery of the features they have in common. Even adults learn and extend object categories better when they are named (e.g., Lupyan, Rakison, & McClelland, 2007). Playing with shape toys that contain alternate shapes may well spur greater understanding of the features of shape categories. This work provides a basis on which to create a straightforward intervention using alternate shapes to test causal impacts on shape knowledge and spatial skills.

7.1. Spatial apps produce many shape name tokens

Prior work on the influence of electronic toys on parent–child interactions indicated that electronic and digital toys can be disruptive to parents and children producing speech during interactions with them (Sosa, 2016; Zosh et al., 2015). Indeed, the outcomes of this study comport with those prior studies and our hypotheses; children heard and produced fewer tokens and types of nearly all kinds when playing with the tablet versus the tangible shapes. The only exception, which is a notable one, is that children in the tablet condition tended to hear markedly more shape tokens when you include the parent and the tablet than the parents produced in either of the tangible toy conditions (an average of 92 tokens vs. 40 and 44). While the tablet itself produced many shape words, it should also be noted, that as in most shape apps thus far available on the market (Resnick et al., 2016), only canonical shapes are pre-

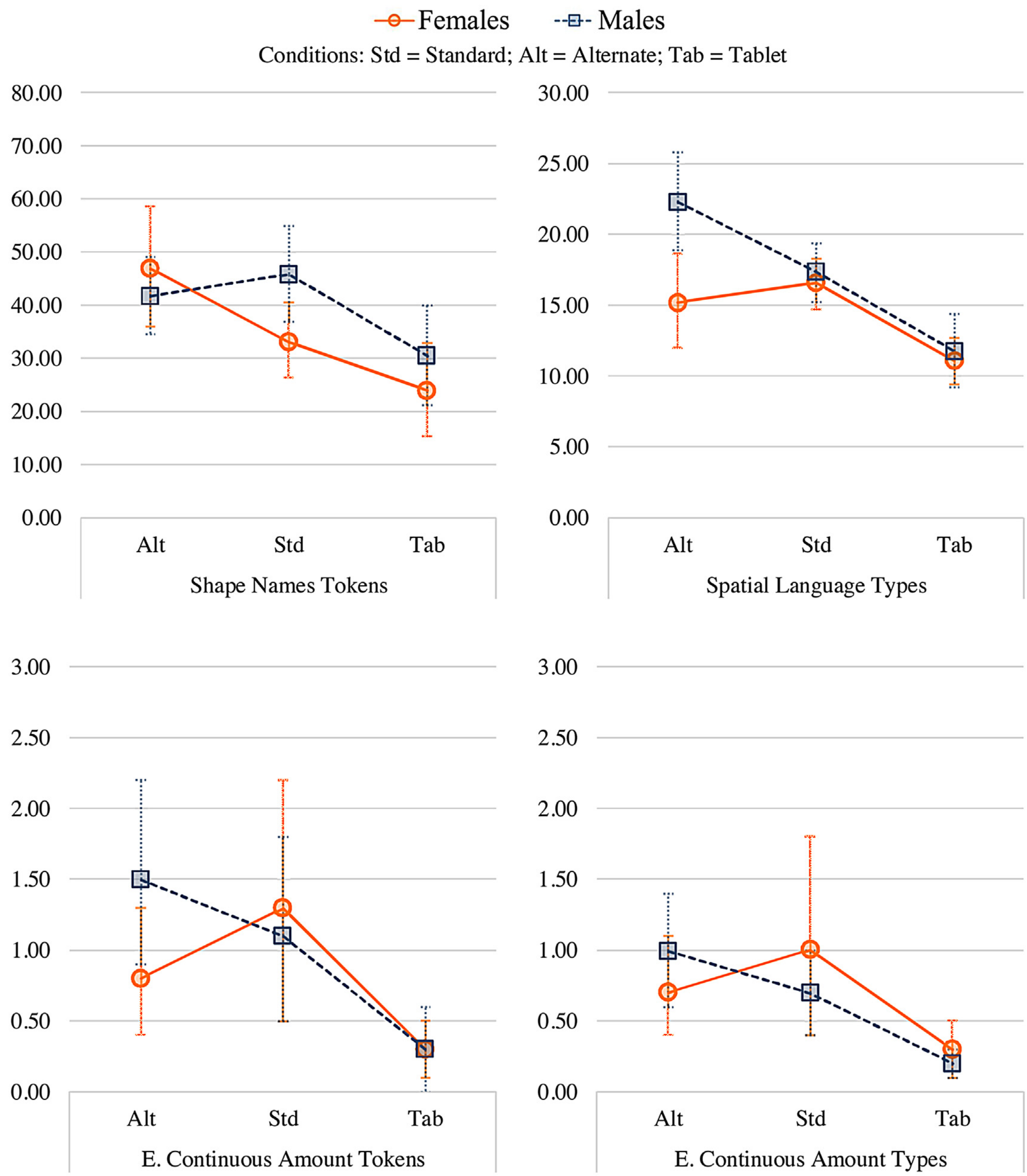

Fig. 3. Bootstrapped means and 95% confidence intervals for Parent Production variables with regressions showing significant effects of sex (Shape Name Tokens), significant interactions of sex and condition (Spatial Language Types), or both (E. Continuous Amount Tokens and Types). Lines represent the sex of the child.

sented. Though this type of "drilling" might be helpful with a highly varied shape set featuring novel examples of categories, repeatedly showing the same set of just a few shapes is not likely to help children key in on the truly defining features of the categories. This approach is analogous to trying to teach the differences between a dog and a cat by repeatedly showing children a black lab and an orange-striped house cat. Does their color matter for being a dog or cat? Size?

Our findings on how children and parents interact with electronic toys are largely consistent with previous research demonstrating that digital formats elicit less language during interactions than traditional, non-electronic versions (Zosh et al., 2015). Such findings present a concern for children whose parents prefer electronic toys or believe that they are highly educational, perhaps due to their classification in the app store as "educational" or the marketing claims made by many companies. Compounding this concern is that lower-income parents are more likely to hold the belief that digital media are highly educational (DeLoache et al., 2010; Hart Research Associates, 2009), which may only exacerbate SES differences in young children.

It is interesting to consider that with the tablet, most of the shape names were produced by the tablet as a result of children interacting with the shapes on screen. Therefore, with only a *very* minor design change, children would have had the opportunity to hear shape names associated with varied visual representations for a category. It is true that there are limitations of tablet technology and differences in the affordances of the devices versus the interactions children can have with humans. For example, tablets cannot easily strike up conversations about shape properties or discuss sophisticated spatial concepts. Indeed the tablet application used here did not produce any spatial words describing the shapes and nearly all of the questions it produced were a prompt to point to a shape ("Where is the [shape name]?"). If you consider that electronic toys are often given to children by parents without a specific intention to be a part of the play situation, these limitations are even more concerning.

Nonetheless, there are many instances in which electronic toys, even given the limitations of current technology, could easily be designed to be more in line with educational principles. And there is evidence that high-quality prompts from electronic media can be effective in learning (e.g., Strouse & Ganea, 2016). Certainly they could be designed to be more encouraging of effective parent–child interactions. Joining other work which spotlights the importance of design in achieving educational goals with electronic media (Clark, Tanner-Smith, & Killingsworth, 2016), the present findings serve to highlight the importance of well-conceived design in eliciting desirable behaviors from parents and children without the need for complicated and expensive interventions.

7.2. Exploratory analyses

7.2.1. Sex differences

Overall, sex differences for children's production were nonexistent in our sample, which is well in line with much of the spatial research on preschool children (Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2017) and prior studies of parent–child interactions in preschoolers (e.g., Zosh et al., 2015). Parents, on the other hand, do produce marginally more shape names with boys and appeared to respond to the alternate shape condition differently depending on the sex of their child. Parents with boys tended to respond to the alternate shape condition in ways that more closely matched our initial hypotheses; they tended to use more spatial language types and tokens (top right panel of Fig. 3). These analyses were exploratory, the p values relatively small, and some other studies looking at parent–child interactions around spatial activities have not found sex differences (Borriello & Liben, 2017), so these effects should be interpreted in that context and with caution.

Nonetheless, there is evidence that parent differences in their use of spatial language with children of different sexes relates to later sex differences in spatial language use (Pruden & Levine, 2017), so additional research is warranted. The apparent patterns observed in this study of shape-based interactions and other research are generally consistent with accounts that sex differences in spatial skills emerge as children age, and that the differences are likely driven by differential exposure to spatial concepts and toys (Baenninger & Newcombe, 1995). It should be mentioned that in dyadic interactions it is often hard to determine whether one speaker is leading or following; an effect for one speaker could create a similar effect for the other. In this instance, it is unlikely that any of the patterns noted in the language parents were using were being driven by what the children were talking about; there is not even a hint of a sex difference in anything the children produced (Table 1 and Table S11.C).

7.2.2. Spatial language subsets

The only spatial language subset analyses that indicated significantly higher production for children in the alternate shape set were the number of deictic tokens. Our interpretation is that their overall increase in spatial language for the alternate condition was due to modest increases across a number of spatial subcategories; the alternate condition had the highest or was tied for the highest spatial language tokens for every subcategory (except D- Orientations and Transformations and H- Pattern, which had means that were effectively 0 for all conditions). Most of these increases were not independently large enough to produce significant regressions given our sample size and length of the transcribed interactions. Parents in the tablet condition produced fewer spatial tokens relative to the standard shape set and significant regressions for C- Location and Direction and E- Continuous Amount indicate that the effect was largest for those subcategories. However, again, the tablet condition tended to have lower means for every category of spatial language relative to one or both of the other conditions.

Low mean counts for many of the subcategories due to the relatively short interactions require caution in interpreting these results. Nonetheless, from these data, it would be hard to conclude that any specific type of spatial language was unilaterally influenced by condition for either parents or children. It would also be hard to conclude that any of the spatial subsets definitely were not influenced by condition. Thus, we consider these outcomes in line with a general effect on spatial language rather than a highly pointed one, though more research is obviously needed.

7.3. Educational implications

These findings on the effects of alternate shapes on parent–child conversation about geometric shapes highlight the importance of toy design on the shape language our children hear. Analogous to the "30 million word gap," Libertus and Golinkoff (2017) have referred to the "The Great Shortchange" as emphasizing the low prevalence of discussion of math concepts. Some parents of toddlers use an average of more than 30 number words every hour (e.g., "Where are the three girls in the picture?̈. Other parents, however, use only one number word every two hours, on average (Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson, 2010). This creates close to a 6000 percent difference in math input at home. The results are even more worrisome when we look at preschool teachers. Some preschool teachers use more than 100 math words per hour, but others use only one – almost a 10,000 percent difference (Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006). Given these differences in children's experience, it is no surprise that some children are better prepared to acquire new math concepts and enter kindergarten with that advantage. Although some children can barely count to 10, others are already doing basic arithmetic. Thus, it is essential to find simple ways to increase the prevalence of spatial and mathematical discussions between parents, children, and educators. Given the recent Common Sense Media findings that mobile device use has tripled among young children (Rideout, 2017), digital prompts that provide scaffolding and quality stimuli that can be used with parents or individually during play are needed. The results of this study suggest that one way to make these improvements is by tweaking shape sets (and other shape toys such as puzzles and blocks) to include more varied shapes and perhaps providing hints to parents about what to talk about as they play.

7.4. Limitations and future directions

The fact that the app did not contain any atypical or alternate versions of the shapes limits our ability to generalize the findings from the two tangible shape conditions to digital materials. Ideally we would have completed a fully crossed design using both standard and alternate shape sets with both the digital and tangible materials. There were, however, no apps available on the market at the time that included an appropriate set of varied and unusual or less-canonical shapes. Comparing from app to tangible shapes is further complicated by the fact that the structure of the game in the app was notably different from how some parents and children chose to play with the tangible shapes. In the real world, outside of a lab setting, children are often handed a tablet to play with independently, without input from parents. As children's technology use continues to become more prevalent, it will be crucial to understand and explore potential features that may improve children's learning from apps. Perhaps future apps will incorporate more varied shapes and rely more heavily on learning science design principles (Hirsh-Pasek et al., 2015).

Additionally, the conditions of the experiment were well controlled to facilitate comparison, but the means for the variables are likely not reflective of what those means might be in a typical, naturalistic situation in the home. Although it is unlikely to have generated the condition differences, given that parents were instructed to teach their children shapes in this more guided context, it seems likely that the means for spatial talk and some of the other variables in this study are higher per unit of time than they would be if they were generated from a free play session in the home, where the materials were simply present. Our sample was also fairly limited in terms of race and socioeconomic status, which likely influences how the data generalize. Almost all parents had college degrees and the sample was predominantly white. Future research is needed with a more diverse sample to examine possible demographic differences (e.g. SES) in parent–child interaction quality and the influences of the condition manipulations. Shape and spatial language use may also be mediated by parent knowledge of shapes and spatial concepts, which may be another compelling area for future research and would be essential in translating some of the current findings into real-world applications.

8. Conclusions

This paper has shown that toy design matters for children hearing and using spatial language that can facilitate their understanding geometric shape properties. Overall, apps and other electronic toys appear to discourage parental and child spatial and shape talk. Until such time that these apps include shape variants and encourage parent–child interaction, this paper suggests that tangible materials tend to elicit more spatial language from children and adults of the type that appear likely to foster shape knowledge and spatial skills. Regardless of the toy type used, variations in the shapes included in shape toy sets may be key to encouraging comparison and discovery of the common features of shapes within the same category. Given the importance of exposing children to shape language for their future spatial competence (e.g., Pruden et al., 2011), these findings have significant implications for toy design and for educational materials.

Author note

This research was supported by the National Institutes of Health Stimulus Grant1RC1HD0634970-01, the National Science Foundation via the Spatial Intelligence and Learning Center (SBE-1041707), and the Institute of Education Sciences through Grant R305A140385.

The authors would like to thank Natalie Brezack for coordinating aspects of the project and data collection, and Josh Pasek for his statistical consultation. We would also like to thank the Child's Play, Learning, and Development Lab at the University of Delaware, especially those who helped with data collection and the creation and execution of the transcription and coding protocols: Sandy Abu El Adas, Kyla Amick, Emma Blackney, Kim Cummings, Jessica Curran, Brendan Czupryna, Laura DiRusso, Emily Horwitz, Paige McHale, Shannon McLaughlin, Adam Rosen, Alexa Rubilotta, Kelly Sassa, Olivia Smith, Hannah Straub, Dunia Tonob, Natalie Wallace, Ori Zaff. We would also like to thank the Spatial Intelligence and Learning Center and their Spatial Network for their consultation at various stages of this project.

Appendix A. Shape interactions – tangible conditions script

The session we are going to tape will have you and your child using wooden shapes. We want you to try to use the shapes to teach your child shape names, the same as you might if you were at home using them. To start off I want to show you the shapes. You'll notice that there are two of each type of shape. For example, there are two circles. [Show the parent the stack of shapes, giving the parents a little bit of time to look at the shapes.]

I'm going to put these shapes in a box, and when we leave the room, you can open the box and start playing. You can use these shapes however you'd like, but here are some ideas for some games you can play!

You can put out three or four of the shapes, and ask your child to find one of them and give it to you, then replace the shape with a different one and play again [experimenter acts out each game quickly and pretty casually]. *You can go through the shapes one-by-one and name them for your child. Another fun game is to pick up a shape and ask your child to find the other one (by name, of course!)*

Hanging on the wall (point to sheet) is a list of the shapes and their names that we want you to try to teach. We have put that there for your reference. It is important that you use the names as they are written up there. For example, please call the moon shape a crescent and not a "moon." We also ask that you not use that sheet to teach (child's name) the shapes. Only use it for your reference.

Whenever (child's name) feels comfortable we are just going leave the room for 10 minutes. During that time we just want you to use the materials we gave you to teach (child's name) as much as you can about shapes. Do you have any questions?

The only other thing I want to add is that our camera is going to be right up here. It would be really helpful to us if you did not sit between the camera and (child's name) so that we can be sure the camera will be able to see what both of you are doing.

Appendix B. Shape interactions – tablet condition script

The session we are going to tape will have you and your child using an iPad app. We want you to try to use the iPad to teach your child shape names, the same as you might if you were at home using one. To start off I want to show you the app and get you a little bit comfortable with using it. [Give the iPad to the parent.]

When you get the iPad out of the box, hit the home button, and it is going to be open to the main menu of the app. [Show them the home key and quickly run through each of the 4 modes (Quizzing, Flashcards, Puzzle, Toy Box). Do not say anything positive or negative about each mode, or indicate that you think one mode is better than any of the others. Please follow the script exactly.]

I'm going to show you each of the games you can choose to use, just so you have an idea of what they are before we start the session. It is up to you and your child what game you want to use.

1. *The first game is called "Quizzing." In this game, a shape comes up, the iPad says the name of the shape and your child has to touch shapes until they get it right.*
2. *The second game is called "Flashcards." In this game, the iPad says the name of the shape and you can swipe left or right to see different shapes.*
3. *The third game is called "Puzzle." In this game your child will have to move each of the shapes into an outline for the shapes.*
4. The last game is called "Toy Box." Please **do not** use this game when playing with your child.

Hanging on the wall (point to sheet) is a list of the shapes and their names that we want you to try to teach. We have put that there for your reference. It is important that you use the names as they are written up there. For example, please call the moon shape a crescent and not a "moon." We also ask that you not use that sheet to teach (child's name) the shapes. Only use it for your reference.

Whenever (child's name) feels comfortable we are just going leave the room for 10 minutes. During that time we just want you to use the *iPad and app we gave you to teach (child's name) as much as you can about shapes. Do you have any questions?*

The only other thing I want to add is that our camera is going to be right up here. It would be really helpful to us if you did not sit between the camera and (child's name) so that we can be sure the camera will be able to see what both of you are doing.

Appendix C. Word categories searched for in the transcripts

Categories below that start with letters are from a spatial language coding manual by Cannon et al. (2007). Our *Shape Names* variable included additional shape terms not listed in their manual and thus we have excluded the letter B from the variable name in tables and statistical reports to be clear that the variable included additional shape terms. Our *Spatial Language* variable is all of the remaining spatial words from their other categories, excluding words counted for *Shape Names*.

Shape names:

B. Shapes – circle, circles, circle's, oval, ovals, oval's, ellipse, ellipses, ellipse's, semicircle, semicircles, semicircle's, triangle, triangles, triangle's, square, squares, square's, rectangle, rectangles, rectangle's, diamond, diamonds, diamond's, pentagon, pentagons, pentagon's, hexagon, hexagons, hexagon's, octagon, octagons, octagon's, parallelogram, parallelograms, parallelogram's, quadrilateral, quadrilaterals, quadrilateral's, rhombus, rhombuss, rhombus's, rhomboid, polygon, sphere, spheres, sphere's, globe, globes, globe's, cone, cones, cone's, cylinder, cylinders, cylinder's, pyramid, pyramids, pyramid's, cube, cubes, cube's, rectangular prism, rectangular prisms, rectangular prism's, shape

Additional shape terms – crescent, crescents, crescent's, heart, hearts, heart's, arrow, arrows, arrow's, clover, clovers, clover's, cross, crosses, cross's, dart, darts, dart's, decagon, decagons, decagon's, egg, eggs, egg's, enneagon, enneagons, enneagon's, heptagon, heptagons, heptagon's, kite, kites, kite's, moon, moons, moon's,nonagon, nonagons, nonagon's, prism, prisms, prism's, quadrafoil, quadrafoils, quadrafoil's, star, stars, star's, trapezium, trapezia, trapezium's, trapezoid, trapezoids, trapezoid's

Spatial language – w/o Shape Names:

A. Spatial dimensions – area, areas, big, bigger, biggest, capacities, capacity, deep, deeper, deepest, depth, depths, emptier, emptiest, empty, enormous, fat, fatter, fattest, full, fuller, fullest, gigantic, height, heights, huge, itsy-bitsy, itty-bitty, large, larger, largest, length, lengths, little, littler, littlest, measure, measurement, measures, narrow, narrower, narrowest, shallow, shallower, shallowest, short, shorter, shortest, skinnier, skinniest, skinny, small, smaller, smallest, tall, taller, tallest, teeny, thick, thicker, thickest, thin, thinner, thinnest, tinier, tiniest, tiny, volume, volumes, wide, wider, widest, width, widths

C. Location and direction – about, above, across, against, ahead, along, among, apart, around, around, around, aside, at, away, back, back, backward, behind, below, beneath, beside, between, beyond, bottom, by, center, close, closer, closest, column, columns, diagonal, direction, directions, distance, distances, down, down, downer, downward, east, eastern, far, forward, from, front, further, head, headed, heading, heads, high, higher, highest, horizontal, horizontally, in, inside, into, join, joined, left, lengthwise, location, locations, low, lower, lowest, middle, near, nearby, nearer, nearest, next to, north, northern, off, on, onto, opposite, out, outside, over, over, parallel, past, path, paths, perpendicular, place, places, position, positions, reverse, reverse, right, route, routes, row, rows, separate, separated, sideways, south, southern, through, throughout, to, together, top, toward, under, underneath, up, upon, upper, upward, vertical, vertically, west, western, with, within

D. Orientation and transformation – flip, orientation, right side up, rotate, rotation, turn, upright, upside down

E. Continuous amount – all, bit, bits, eight, eights, eigth, eigths, enough, equal, fifth, fifths, fraction, fractions, fragment, fragments, half, halfs, less, little, lot, more, much, ninth, ninths, none, part, parts, piece, pieces, portion, portions, quarter, quarters, same, section, sections, segment, segments, seventh, sevenths, sixth, sixths, some, space, tenth, tenths, third, thirds, whole, wholes

F. Deictics – anywhere, everywhere, here, nowhere, somewhere, there, where, wherever

G. Spatial features and properties – angle, angles, arc, arcs, axes, axis, bend, bended, bends, bendy, bent, border, bordered, borders, bump, bumped, bumps, bumpy, circular, conical, corner, corners, curve, curved, curves, curvey, cylindric, cylindrical, diagonal, diagonals, edge, edged, edges, elliptical, face, faces, flat, flatter, flattest, horizontal, line, lined, lines, lump, lumps, lumpy, parallel, perpendicular, plane, planes, point, pointed, points, rectangular, round, rounded, rounder, roundest, sector, sectors, shaped, side, sided, sides, spheric, spherical, surface, surfaces, symmetric, symmetrical, symmetry, triangular, vertical, wave, waves, wavey

H. Pattern – after, before, decrease, decreased, decreases, decreasing, design, designs, first, increase, increased, increases, increasing, last, next, order, orders, pattern, patterns, repeat, repeated, repeating, repeats, repetition, sequence, sequences

Number words:

One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty, fifty, hundred.

Appendix D. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ecresq.2018.03.015.

References

Baenninger, M., & Newcombe, N. S. (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learning and Individual Differences*, *7*(4), 363–379. http://dx.doi.org/10.1016/1041-6080(95)90007-1

Borriello, G. A., & Liben, L. S. (2017). Encouraging maternal guidance of preschoolers' spatial thinking during block play. *Child Development*, http://dx.doi.org/10.1111/cdev.12779

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65*(1), 14–21. http://dx.doi.org/10.1037/h0041727

Cannon, J., Levine, S. C., & Huttenlocher, J. (2007). *A system for analyzing children and caregivers' language about space in structured and unstructured contexts. Spatial Intelligence and Learning Center (SILC) technical report.*

Carmichael, C. A., & Hayes, B. K. (2001). Prior knowledge and exemplar encoding in children's concept acquisition. *Child Development*, *72*(4), 1071–1090. http://dx.doi.org/10.1111/1467-8624.00335

Cartmill, E., Pruden, S. M., Levine, S. C., & Goldin-Meadow, S. (2010). The role of parent gesture in children's spatial language development. In K. Franich, K. M. Iserman, & Keil (Eds.), *Proceedings of the 34th annual Boston University conference on language development* (pp. 70–77). Somerville, MA: Cascadilla Press. Retrieved from http://wew.spatiallearning.org/publicationspdfs/Cartmill%20Pruden%20Levine%20%20Goldin-Meadow%20(in%20press).pdf

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79–122. Retrieved from http://journals.sagepub.com/doi/abs/10.3102/0034654315582065

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 420–464). New York, NY: Macmillan Publishing.

Clements, D. H., Sarama, J. H., & Liu, X. H. (2008). Development of a measure of early mathematics achievement using the Rasch model: The Research-based early maths assessment. *Educational Psychology*, *28*(4), 457–482. http://dx.doi.org/10.1080/01443410701777272

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*(2), 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

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., & O'Doherty, K. (2010). Do babies learn from baby media? *Psychological Science*, 21(11), 1570–1574. Retrieved from http://pss.sagepub.com/content/21/11/1570.short

Ferrara, K., Hirsh-Pasek, K., Newcombe, N. S., Golinkoff, R. M., & Lam, W. S. (2011). Block talk: Spatial language during block play. *Mind, Brain, and Education*, *5*(3), 143–151. http://dx.doi.org/10.1111/j.1751-228X.2011.01122.x

Fisher, K. R., Hirsh-Pasek, K., Newcombe, N. S., & Golinkoff, R. M. (2013). Taking shape: Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development*, *84*(6), 1872–1878. http://dx.doi.org/10.1111/cdev.12091

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends in Cognitive Sciences*, *18*(10), 536–542. http://dx.doi.org/10.1016/j.tics.2014.05.011

Gentner, D. (2016). Language as cognitive tool kit: How language supports relational thought. *American Psychologist*, *71*(8), 650–657. http://dx.doi.org/10.1037/amp0000082

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development*, *14*(4), 487–513. Retrieved from http://www.sciencedirect.com/science/article/pii/S0885201499000167

Hart Research Associates. (2009, November). *Parenting infants and toddlers today research findings: Based on a survey of parents of children ages birth to three years old. Conducted on behalf of zero to three*. Retrieved from http://www.zerotothree.org/about-us/funded-projects/parenting-resources/final_survey_report_3-11-2010.pdf

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89.

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in 'educational' apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3–34. Retrieved from http://psi.sagepub.com/content/16/1/3.abstract

Huttenlocher, J., & Levine, S. C. (1990). *Primary test of cognitive skills*. CTB/McGraw-Hill.

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher math talk. *Developmental Psychology*, *42*(1), 59–69. http://dx.doi.org/10.1037/0012-1649.42.1.59

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*(4), 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*(2), 530–542. http://dx.doi.org/10.1037/a0025913

Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology*, *46*(5), 1309–1319. http://dx.doi.org/10.1037/a0019671

Libertus, M. E., & Golinkoff, R. M. (2017, January). *Commentary: good math skills begin at home*. Philadelphia Inquirer. Retrieved from http://www.philly.com/philly/opinion/20170102_Commentary__Good_math_skills_begin_at_home.html

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

Lombardi, C. M., Casey, B. M., Thomson, D., Nguyen, H. N., & Dearing, E. (2017). Maternal support of young children's planning and spatial concept learning as predictors of later math (and reading) achievement. *Early Childhood Research Quarterly*, *41*, 114–125. http://dx.doi.org/10.1016/j.ecresq.2017.07.004

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science*, *18*(12), 1077–1083. http://dx.doi.org/10.1111/j. 1467–9280.2007.02028.x

MacWhinney, B. (2000). (3rd ed.). *The CHILDES project: Tools for analyzing talk* (Vol. 2) Mahwah, NJ: Lawrence Erlbaum Associates.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA, US: MIT Press.

Nazareth, A., Herrera, A., & Pruden, S. M. (2013). Explaining sex differences in mental rotation: Role of spatial activity experience. *Cognitive Processing*, *14*(2), 201–204. http://dx.doi.org/10.1007/s10339-013-0542-8

Newcombe, N. S. (2017). Harnessing spatial thinking to support stem learning. *OECD education working papers*, *161*, 1–51. http://dx.doi.org/10.1787/7d5dcae6-en

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*. U.S. Department of Health and Human Services. Retrieved from http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/eecd/Assessment/Child%20Outcomes/HS_Revised_Child_Outcomes_Framework(rev-Sept2011).pdf

Polinsky, N., Perez, J., Grehl, M., & McCrink, K. (2017). Encouraging spatial talk: Using children's museums to bolster spatial reasoning. *Mind, Brain, and Education*, *11*(3), 144–152. http://dx.doi.org/10.1111/mbe.12145

Pruden, S. M., & Levine, S. C. (2017). Parents' spatial language mediates a sex difference in preschoolers' spatial language use. *Psychological Science*, 1583–1596. http://dx.doi.org/10.1177/0956797617711968

Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science*, *14*(6), 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.

Ramani, G. B., Rowe, M. L., Eason, S. H., & Leech, K. A. (2015). Math talk during informal learning activities in Head Start families. *Cognitive Development*, *35*, 15–33. http://dx.doi.org/10.1016/j.cogdev.2014.11.002

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

Rideout, V. J. (2017). *The common sense census: Media use by kids age zero to eight*. San Francisco, CA: Common Sense Media.

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*(1), 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*(1), 181–189. http://dx.doi.org/10.1016/j.ecresq.2004.01.014

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*(3), 662–674. http://dx.doi.org/10.1016/j.cognition.2008.05.007

Satlow, E., & Newcombe, N. S. (1998). When is a triangle not a triangle? Young children's developing concepts of geometric shape. *Cognitive Development*, *13*(4), 547–559. http://dx.doi.org/10.1016/S0885-2014(98)90006-5

Seo, K.-H., & Ginsburg, H. P. (2004). What is developmentally appropriate in early childhood mathematics education? Lessons from new research. In D. H. Clements, J. H. Sarama, & A. M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 91–104). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*(3), 545–560. http://dx.doi.org/10.1037/a0014239

Sosa, A. V. (2016). Association of the type of toy used during play with the quantity and quality of parent–infant communication. *JAMA Pediatrics*, 1–6. http://dx.doi.org/10.1001/jamapediatrics.2015.3753

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, *19*(1), 99–120. http://dx.doi.org/10.1016/j.ecresq.2004.01.002

Strouse, G. A., & Ganea, P. A. (2016). Are prompts provided by electronic books as effective for teaching preschoolers a biological concept as those provided by adults? *Early Education and Development*, 1–15. http://dx.doi.org/10.1080/10409289.2016.1210457

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology*, *49*(12), 2368–2382. http://dx.doi.org/10.1037/a0032463

Szechter, L. E., & Liben, L. S. (2004). Parental guidance in preschoolers' understanding of spatial-graphic representations. *Child Development*, *75*(3), 869–885. http://dx.doi.org/10.1111/j.1467-8624.2004.00711.x

Terlecki, M. S., & Newcombe, N. S. (2005). How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles*, *53*(5-6), 433–441. http://dx.doi.org/10.1007/s11199-005-6765-0

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*, *1*, 142–161. http://dx.doi.org/10.1080/15248372.2015.1016610

Verdine, B. N., Bunger, A., Athanasopoulou, A., Golinkoff, R. M., & Hirsh-Pasek, K. (2017). Shape up: An eye-tracking study of preschoolers' geometry knowledge and spatial development. *Developmental Psychology*, *53*(10), 1869–1880. http://dx.doi.org/10.1037/dev0000384

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), 1–150. http://dx.doi.org/10.1111/mono.12263

Waxman, S. R., & Hall, D. G. (1993). The development of a linkage between count nouns and object categories: Evidence from fifteen- to twenty-one-month-old infants. *Child Development*, *64*(4), 1224. http://dx.doi.org/10.2307/1131336