

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

General Mechanisms Underlying Language and Spatial Cognitive Development

Permalink

<https://escholarship.org/uc/item/5s7380dz>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors

Miller, Hilary E.

Simmering, Vanessa R.

Publication Date

2016

Peer reviewed

General Mechanisms Underlying Language and Spatial Cognitive Development

Hilary E. Miller (hemiller@wisc.edu) and Vanessa R. Simmering (simmering@wisc.edu)

Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Abstract

Previous research showed that children's spatial language production predicts their spatial skills, but the mechanisms underlying this relation remain a source of debate. This study examined whether 4-year-olds' spatial skills were predicted by their attention to task-relevant information—in tasks that emphasize either memory or language—above and beyond their spatial word production. Children completed three types of tasks: (1) a memory task assessing attention to task-relevant color, size, and location cues; (2) a production task assessing adaptive use of language to describe scenes, varying in color, size, and location; and (3) spatial tasks. After controlling for age, gender, and vocabulary, children's spatial skills were significantly predicted by their memory for task-relevant cues, above and beyond their task-related language production and adaptive use of language. These findings suggest that attending to relevant information is a process supporting spatial skill acquisition and underlies the relation between language and spatial cognition.

Keywords: spatial cognition; short-term memory; language production; cognitive development

Introduction

Spatial skills are fundamental cognitive abilities that support basic behaviors such as perceiving and remembering locations, navigating, and relational reasoning. Additionally, spatial skills predict performance in math and science and entry into STEM fields (Verdine et al., 2014; Wai, Lubinski, & Benbow, 2009). As individual differences in spatial skills arise early in development (e.g., Verdine et al., 2014), it is important to understand the factors that contribute to the development of spatial skills during early childhood in order to devise effective interventions to promote spatial cognition.

Spatial language has been identified as one factor that predicts children's spatial abilities. Multiple studies have identified relations between language and spatial cognition across a variety of spatial tasks (e.g., Dessalegn & Landau, 2008; Pruden, Levine, & Huttenlocher, 2011). Despite the proliferation of research, the causal nature of the link between language and spatial cognition remains unclear. It is possible that language directly causes changes in spatial abilities, spatial abilities cause changes in language, language indirectly relates to spatial cognition through general mechanisms such as attention, or that multiple causal factors interact bi-directionally to shape spatial skills and language. The current study investigated whether a more general mechanism of attention to task-relevant cues might underlie the relation between language and spatial cognition.

Previous research has shown correlations between children's spatial word production and their spatial performance. For example, children who produced particular

spatial terms such as “left” and “right” (Hermer-Vazquez, Moffet, & Munkholm, 2001) and “middle” (Simms & Gentner, 2008) tended to perform better on spatial tasks that involve these dimensions. Also, the total number of spatial words that children spontaneously produced during free play predicted their spatial performance (Pruden et al., 2011). These effects are thought to be specific to spatial words, as the relation between spatial word production and children's performance on spatial tasks held even after controlling for spatial word comprehension, IQ, (Hermer-Vazquez et al., 2001) and general receptive vocabulary (Pruden et al., 2011).

Further evidence that language facilitates spatial development comes from paradigms in which children were provided with verbal cues specifying spatial relations among objects or features (e.g., “top”, “left”) before or during spatial tasks. The results across these studies showed that spatial language cues enhanced children's spatial task performance (e.g., Dessalegn & Landau, 2008; Miller, Patterson, & Simmering, 2016). These findings have been interpreted as evidence that language helps children verbally encode task-relevant spatial information and once children have acquired relevant spatial words they can use language to facilitate their performance (e.g., Hermer-Vazquez et al., 2001; Loewenstein & Gentner, 2005; Pruden et al., 2011).

Other studies showed that such effects were not specific to spatial language, but arose more generally through highlighting task-relevant cues (e.g., Shusterman, Lee, & Spelke, 2011). Specifically, Shusterman et al. (2011) showed that providing children with non-spatial language (e.g., “the red wall can help you find the sticker”) that highlighted the utility of particular cues improved their spatial performance. This effect was specific to task-relevant language, as similar but task-irrelevant language did not improve performance (e.g., “look at the pretty red wall”).

The range of results across conditions with both spatial and non-spatial words suggests that language can draw children's attention to relevant cues in a spatial task. However, these findings raise questions about whether the relation between language and spatial cognition arises from children's spatial word acquisition. Is it the case that once children can produce relevant language (e.g., spatial words) they use their language knowledge to verbally encode task-relevant information or is it the case that children need to be selective in the types of cues they encode? It is possible that children can produce task-relevant language on their own, but may not use it in the context of spatial tasks. Being provided with language could help children encode relevant information, but their abilities to produce words on their own may not be enough to direct children to encode relevant cues.

Miller, Vlach, and Simmering (in press) tested these two

alternative accounts and showed that the relation between language and spatial cognition does not arise solely from children’s abilities to produce spatial words. Rather, it reflects children’s abilities to adaptively use language. In this study, 4-year-olds completed a spatial scene description task that assessed their use of task-relevant language. In the task, children described the location of a mouse in a spatial scene (Figure 1). There were three possible types of cues (color, size, or location) that children could use to describe the location of the mouse; across trials the relevance of the color and size cues varied such that on some trials, color and/or size was not a relevant cue for describing the mouse’s location (cf. Figure 1A vs 1B or 1C). Results showed that children who used language more adaptively (i.e., provided more relevant than irrelevant cues) performed better on the spatial tasks, even when controlling for their age, gender, vocabulary (both general receptive and spatial productive), and quantity of task-related language produced (including spatial words).

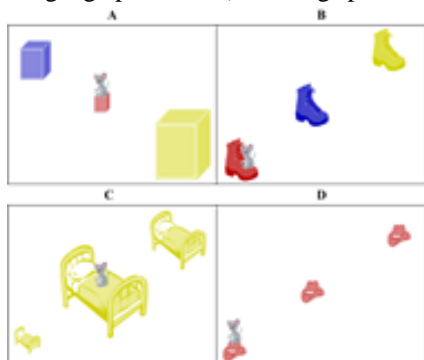


Figure 1: Sample trials of the spatial scene description task with varying numbers of distinctive cues: A) 3 cues = color, size, and/or location, B) 2 cues = color and/or location, C) 2 cues = size and/or location, D) 1 cue = location only.

Miller et al.’s (in press) results suggest that using language in task-relevant ways relates to spatial performance, but it is unclear whether language is the causal factor underlying this link or whether a third factor—not specific to language or spatial skills—contributes to both. Miller et al. hypothesized that children’s attention to task-relevant cues supports both language and spatial skills. Some support for this hypothesis comes from research showing that children’s spatial task performance can be facilitated by non-spatial language (e.g., Shusterman et al., 2011) and/or non-verbal cues that direct attention to task-relevant features (i.e., by making them more stable or salient; Learmonth, Nadel, & Newcombe, 2002; Learmonth, Newcombe, & Huttenlocher, 2001).

The current study provides a further test of the hypothesis that the link between spatial skills and language arises from a third common factor: the ability to direct attention to task-relevant cues. We tested this by comparing children’s spatial task performance to not only their spatial scene descriptions (as in Miller et al., in press), but also a memory task assessing attention to relevant spatial and non-spatial cues (see Figure 2 below). In the memory task, children viewed the scenes

from the description task and after a brief delay had to choose which of three pictures matched the memory array based on color, size, or location of the target referent object. We hypothesized that children’s memory for task-relevant information in this task would predict their spatial skills above and beyond the factors previously shown to relate to children’s spatial cognition: age, gender, PPVT-IV score, task-related production, and adaptation score (Levine, Huttenlocher, Taylor, & Langrock, 1999; Miller et al., in press; Pruden et al., 2011; Voyer, Voyer, & Bryden, 1995).

Method

Participants

Sixty-nine 4-year-olds ($M = 4.52$, $SD = 0.39$ years, 31 females) participated in the study. An additional 13 children participated but were excluded due to: incomplete data (4), experimenter error (1), technical problems (2), parental interference (1), non-compliance (1), and not talking during the production task or insufficient vocabulary (3 and 1, respectively, described further below). Participants were recruited from a database compiled by a university research center. Participants received small prizes for participation.

Design and Procedures

Children were tested individually in the spatial scene description task (description task), the spatial scene memory task (memory task), three spatial tasks, and the Peabody Picture Vocabulary Test (PPVT) IV. Tasks were presented in a quasi-random order such that the first, third, and fifth tasks were always spatial tasks and the second and fourth tasks were the description and memory tasks, and the final task was the PPVT-IV. The order of the spatial tasks and the order of the description and memory task were counter-balanced across participants. Children received small prizes between tasks. Caregivers completed a productive spatial vocabulary checklist. Sessions lasted approximately 45 minutes each.

Spatial Scene Description Task Before the description task, children completed a warm-up task to help them understand the task and to feel comfortable talking aloud. In the warm-up task, children viewed 10 PowerPoint slides showing one familiar object on each slide and were asked to describe what they saw to a stuffed animal who was not looking at the screen (“Tell Bucky what you see”). The experimenter discouraged children from pointing during the warm-up and description tasks by instructing them to sit on their hands.

The description task tested children’s abilities to disambiguate the target object’s location (i.e., mouse) relative to a referent object. The task was conducted in PowerPoint. Each trial showed a picture of a spatial scene in the center of the slide. Each scene included three referent objects (e.g., beds) distributed diagonally across the screen (see Figure 1 above); the diagonal orientation (top-left to bottom-right vs. top-right to bottom-left) varied randomly across trials. The mouse was in a support relation to one referent object (target

referent object). On each trial, the child was asked to “Tell Bucky where the mouse is” to a stuffed animal not facing the screen. The trials varied in the number of relevant cues for describing the mouse’s location such that the child could use 3 cues (color, size, and location, Figure 1A), two cues (color and location, size and location; Figure 1B and 1C), or one cue (location, Figure 1D) to describe the mouse’s location.

The description task included 24 trials presented in one of four randomized orders. There were 6 trials per cue type, with 2 trials for each possible mouse’s location (i.e., front, middle, back referent; see Figure 1). The diagonal alignment of the referent objects allowed children to use a range of different spatial words to describe the referent object’s location (e.g., front, middle, back, first, last, left, right, top, center, bottom).

Spatial Scene Memory Task The memory task was modeled after the description task and tested children’s memory for the features associated with the mouse’s location. The PowerPoint slides from the description task were used in this task (in the same trial order) as memory arrays. Following the memory array presentation, there was a 1 s delay with a blank screen, and then the test array was presented. Test arrays probed memory for one dimension (color, size, or location) of the target referent object (Figure 2). Test arrays included three pictures along the bottom of the screen; these pictures showed only one referent object, and children were instructed to pick the picture that exactly matched the referent (e.g., box, ball, couch) the mouse was on in the memory array. Foil pictures varied along the probe dimension, for example, if the probe was color the referent objects in the foil scenes would be the same size and location as the target referent, but a different color (see Figure 2A). For each cue type (e.g., 2 cue-color and location, Figure 2C) of the memory array, children were presented twice with each probe type. The probe dimensions were randomized across trials and children did not know which probe would be tested in advance.

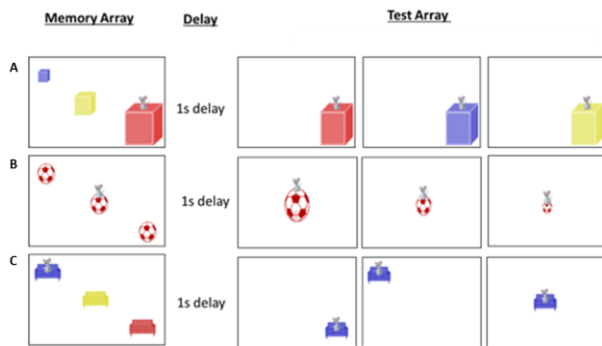


Figure 2: Sample trials of the spatial scene memory task: A) 3-cue memory array, color probe; B) 1-cue memory array, size probe; C) 2-cue memory array, location probe.

Spatial Tasks Children participated in short versions of three spatial tasks (shown in Figure 3): Spatial Analogies Task (Levine et al., 1999; Pruden et al., 2011), Children’s Mental Transformation Task (Huttenlocher & Levine, 1990; Pruden

et al., 2011), and Feature Binding Task (Dessalegn & Landau, 2008). For the Spatial Analogies and Mental Transformation Task, we adapted the short version from Pruden et al. (2011) and for the Feature Binding Task, we used half the number of trials as in Dessalegn & Landau (2008). In the Mental Transformation Task (10 trials; Figure 3A), children saw two pieces of a shape and selected which of four shapes the two pieces would make if combined. In the Spatial Analogies Task (13 trials; Figure 3B), children viewed a picture depicting two objects in a spatial relation and chose which of four other pictures shared that relation. In the Feature Binding Task (12 trials; Figure 3C), children saw a square with two different colors on opposite sides and chose a matching figure out of three options following a 1 s delay.

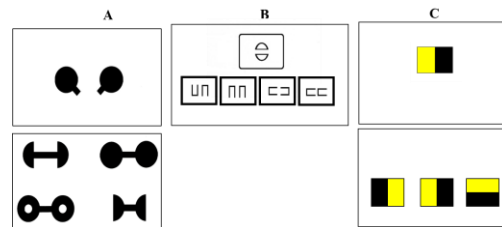


Figure 3: Sample trials of the spatial tasks: A) Mental Transformations; B) Spatial Analogies; C) Feature Binding.

Vocabulary assessments. The PPVT-IV measures receptive vocabulary and involved the children pointing to one of four pictures depicting the target word. The spatial vocabulary checklist measured productive spatial vocabulary and included 80 words taken from both the MCDI: Words and Sentences (Fenson et al., 1994) and from a spatial word coding manual. Caregivers indicated words they have heard their child produce (Cannon, Levine, & Huttenlocher, 2007)

Coding and measurement

Spatial scene description task Each session was transcribed by two research assistants. Transcribers were reliable on 89% of trials and discrepancies were resolved through a third research assistant blind to the first two transcripts. Final transcripts were coded by two different research assistants blind to the study hypotheses (mean 98% reliability across dimensions coded), and disagreements were resolved by the first author. Coders scored the number of times children mentioned color or spatial terms, and whether they used the terms correctly. For spatial terms, we separately categorized both size (e.g., small, medium) and location terms (terms referring to the referent’s location in the scene, e.g., *top* chair) from the other types of spatial terms mentioned.

For the regression analysis, we calculated three measures from the coding as in Miller et al. (in press). For task-related production, we tabulated per trial (1) the quantity of non-spatial terms used by averaging the number of color terms produced, and (2) the quantity of spatial terms used by averaging the number of spatial terms (including but not limited to size and location terms). These variables were

created to ensure that any effects found with children's adaptive use of language did not result solely from the number of potentially-relevant words children could produce.

For (3) the adaptation score, we calculated how often each child produced relevant versus irrelevant cues during the description task. On each trial, a child's response was scored for the three cue types (color, size, and location) as a 0 or 1 depending on whether they produced any description of those cues. Multiple descriptions of the same cue type were only counted as one (i.e., "front corner" and "front" would both count as 1 for location cues on that trial). To account for the relevance of the cues, color and size terms were coded as negative if produced on trials when the terms could not differentiate the referents. Specifically, color was coded as negative on 2 cue (size and location) and 1 cue trials (Figure 1C and 1D), and size was coded as negative on 2 cue (color and location) and 1 cue trials (Figure 1B and 1D). On all other trials color and size cues were coded as positive, and location cues were always coded as positive because location was relevant on all trial types. Positive adaptation scores reflected children providing more relevant than irrelevant cues and negative adaptation scores reflected children providing more irrelevant than relevant cues. Scores closer to 0 reflected performance that did not differ by trial type; for example, a child who mentioned only color terms on every trial would receive a score of 0. Scores farther from 0 reflected both the number of cues mentioned and the cue relevance. Thus, the adaptation score was sensitive to the context in which children provided different types of cues.

To ensure that children had sufficient language knowledge to perform the task, we made two types of exclusions. We excluded trials for which the child's caregiver indicated on the checklist that their child did not produce any of the relevant location words (e.g., trials with the target on the middle object if the checklist indicated the child did not produce *middle* or *center*) and the child did not produce these words during the task. We also excluded trials from the description task if the child used an incorrect color, size, or location term (e.g., saying "top", when the object was on the bottom of the display). This resulted in the exclusion of 121 trials from 25 children's data (8% of total trials).

Spatial scene memory task and spatial tasks The experimenter marked responses on a session sheet during the study. Videos were used to code for reliability; 18 participants' sessions per task (26%, with different participants chosen for each task) were checked by a second research assistant, resulting in 99% agreement with disagreements resolved by a third research assistant. We created a spatial composite score by calculating the mean proportion correct in the three spatial tasks after normalizing for different chance levels (see Figure 3 above). We normalized scores by taking each child's score minus chance, then dividing by one minus chance, resulting in scores with 0 as chance and 1 as perfect performance.

For the memory task, we calculated a memory composite score for each child as their overall proportion of correct

responses. We also grouped the trial types based on whether the probed dimension varied in the memory array (i.e., Figure 2A shows a trial where color was probed when it varied, and Figure 2B shows a trial where size was probed when it did not vary) to parallel the adaptation score from the description task. As in the description task, the memory arrays varied as to whether color and size cues differentiated the target referent object. However, unlike the description task, all cue dimensions were probed across trials and thus were potentially relevant. As such, instead of taking a difference score of relevant vs. irrelevant cues as in the adaptation scores, we created separate variables based on whether the probed dimension differentiated the target referent object.

We created three variables for this analysis: color/size (C/S) differentiated, C/S undifferentiated, and location differentiated. The C/S differentiated variable was calculated as children's mean proportion correct on color-probe trials on which color varied in the memory array (e.g., Figure 2A) and size-probe trials on which size varied in the memory array. The C/S undifferentiated variable was calculated as children's mean proportion correct on color-probe trials on which color did not vary in the memory array and size-probe trials on which size did not vary in the memory array (e.g., Figure 2B). The location differentiated variable was calculated as children's mean proportion correct on all location-probe trials (e.g., Figure 2C), as location always differentiated the referent objects. Although the location differentiated and C/S differentiated variables were similar (i.e., both include probe dimensions that varied in the memory arrays), we chose to calculate them separately to have equal number of trials (8) included for each variable.

Vocabulary assessments The experimenter marked children's responses while administering the PPVT-IV and terminated testing when the child responded incorrectly on eight or more trials within a 12-trial block (following the standardized instructions). Children's standardized scores were calculated offline using established norms (Dunn & Dunn, 2007). Productive spatial vocabulary was calculated as the proportion of words caregivers checked.

Results

This study investigated the relation between children's spatial skills and their attention to task-relevant information in both memory and language tasks. We hypothesized that individual differences in children's memory for relevant cues would predict their spatial composite score above and beyond demographics, vocabulary, quantity of task-related language production, and adaptive use of language. One participant's data was removed for being an outlier in the regression model (final $N = 68$). As a preliminary analysis, we tested whether there were differences in children's spatial composite scores based on whether they received the description or memory task first, but found no significant difference ($p = .926$) and thus excluded this variable from further analyses.

To evaluate our hypothesis, we conducted a regression analysis with two steps. The first step was conducted as a

replication of Miller et al. (in press) to test whether children’s adaptive use of language predicted their spatial task performance. The second step evaluated our hypothesis by testing whether children’s memory for task-relevant information predicted their spatial skills above and beyond their adaptive use of language. For the first step, children’s spatial composite scores were regressed on demographics (age, gender), vocabulary (PPVT-IV, productive spatial vocabulary), quantity of task related language production (quantity of non-spatial and spatial terms used), and adaptation score. We found that children’s adaptation score significantly predicted their spatial composite score ($t_{60} = 2.21$, $p = .031$, $\Delta R^2 = .045$), replicating previous findings (Miller et al., in press). For the second step, we added each child’s memory composite score to the model. We found that proportion correct on the memory task significantly predicted children’s spatial composite score ($t_{59} = 2.43$, $p = .018$, $\Delta R^2 = .050$) as shown in Figure 4A, and the variance accounted for by the adaptation score became marginal ($p = .090$)¹.

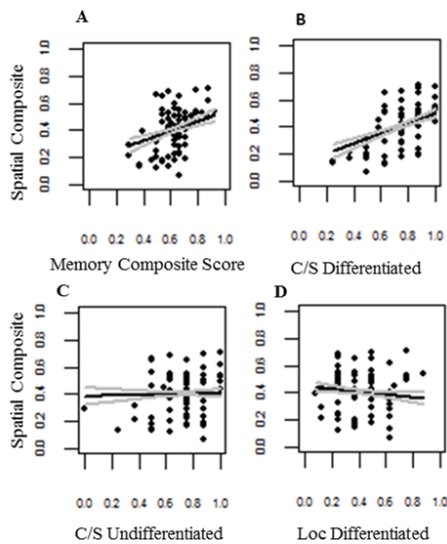


Figure 4: Relation between children’s spatial composite score and A) memory composite score, performance on B) C/S differentiated trials, C) C/S undifferentiated trials, and D) location (Loc) differentiated trials. Error bars represent ± 1 S.E. for point estimates from the regression models.

To test whether performance across the trial types differentially predicted spatial performance, children’s spatial composite scores were regressed on C/S differentiated, C/S undifferentiated, and location differentiated (controlling for demographic, vocabulary, task-related production, and adaptation score variables as before), shown in Figures 4A-4C, respectively. Performance on C/S differentiated trials significantly predicted children’s spatial composite scores ($t_{57} = 3.77$, $p < .001$, $\Delta R^2 = .101$) but

performance on C/S undifferentiated and location differentiated trials did not account for significant proportions of the variance in spatial composite score ($p = .777$, $p = .265$, respectively). The results are consistent with the hypothesis that attention to task-relevant cues is a general mechanism supporting spatial cognition and language use.

Discussion

This study tested whether children’s attention to task-relevant cues in a memory task was more predictive of their spatial skills than their use of task-relevant cues in a production task. We posited that 4-year-olds’ memory for task-relevant cues would predict their spatial composite score above and beyond demographics (age, gender), vocabulary (PPVT-score, productive spatial vocabulary), task-related production (quantity of non-spatial and spatial terms used) and adaptation score. Before adding the memory score, the model showed that adaptive language use predicted children’s spatial composite scores after controlling for the other variables, replicating results that adaptive use of language—both spatial and non-spatial—is more predictive of spatial skills than spatial word production (Miller et al., in press).

Adding the memory composite score to the model showed that memory for task-relevant cues predicted children’s spatial skills above and beyond all factors included in the model. The amount of variance accounted for by the adaptation score became only marginally significant when the memory composite score was added, indicating overlap in the contributions these measures make toward accounting for variance in children’s spatial skills. Overall, these findings suggest that children’s attention to cues in a spatial scene is related to their spatial performance, and could potentially account for much of the variance previously thought to be related to language. However, it is important to note that these findings do not exclude the possibility that other factors such as encoding or memory strategies may explain some of the variance in the memory task.

When we grouped memory task performance in a similar way to the description task, we found that only children’s memory for color and size cues that differentiated the target referent objects in the memory array, when these dimensions were probed, significantly predicted spatial skills. These results are consistent with our hypothesis that attention to task-relevant cues underlies the relation between spatial skills and language. Children who correctly recognized the target referent object’s color or size cues when these cues were differentiated in the memory array likely performed better on the spatial tasks because they are better at attending to cues that distinguish relevant information in the spatial tasks. This ability would also support more relevant language production in the spatial scene description task, leading to the shared variance across these tasks found in our results. Performance on the color and size undifferentiated trials likely did not

¹ The qualitative pattern of results remains the same with the outlier data included. The only quantitative difference is that the adaptation score is marginal in the first step ($p = .089$) and non-

significant in the second step ($p = .225$) with the outlier data included.

predict spatial performance because in the spatial tasks features that do not differentiate among cues would not be relevant for solving the task. Thus, whether or not children attended to non-differentiated cues should not influence their performance as long as they attend to the relevant/differentiated cues. Although location cues always differentiated the target referent object, performance on location probes likely did not predict children's spatial skills due to poor performance ($M = .42$, $SD = .49$; see Figure 4D). Children rarely produced location terms in the spatial scene description task (5% of trials; similar to levels reported by Miller et al., in press), indicating that they may not have attended to location cues in the description or memory tasks. This is interesting because it suggests that attention to locations is not strongly related to spatial skills that rely on transforming or comparing relations among objects.

The current study provides novel insights into the mechanisms underlying the relations between language and spatial cognition. Of the potential causal models laid out in the introduction, our results suggest that a third factor connects spatial language and spatial skills. Further studies will be needed to test potential causal directions among the various factors that support developmental improvements in spatial skills and language use. Our research highlights the importance of considering the multiple cognitive processes that support spatial performance and provides insight into basic cognitive factors that may connect spatial skills to cognitive processes that support achievement in STEM domains

Acknowledgments

Thanks to the families who participated in this research, as well as to the research assistants who aided in data collection. Participant recruitment was partially funded by the Waisman Intellectual and Developmental Disabilities Research Center grant through NICHD (P30HD03352). Subsets of the data were presented at the 9th Biennial Meeting of the Cognitive Development Society and the 56th Annual Meeting of the Psychonomic Society.

References

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.

Dessalegn, B., & Landau, B. (2008). More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological Science*, *19*, 189–195.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test, fourth edition: Manual*. Minneapolis, MN: Pearson Assessments.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in early communicative development. *Monographs of*

the Society for Research in Child Development, *i*–185. <http://doi.org/10.2307/1166093>

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition*, *79*, 263–299.

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

Learmonth, A. E., Nadel, L., & Newcombe, N. S. (2002). Children's use of landmarks: Implications for modularity theory. *Psychological Science*, *13*, 337–341.

Learmonth, A. E., Newcombe, N. S., & Huttenlocher, J. (2001). Toddlers' use of metric information and landmarks to reorient. *Journal of Experimental Child Psychology*, *80*, 225–244.

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

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*, 315–353.

Miller, H. E., Patterson, R., & Simmering, V. R. (2016). The role of language in children's use of spatial relations to remember locations. *Cognition*, *150*, 170–180.

Miller, H. E., Vlach, H. A., & Simmering, V. R. (in press). Producing spatial words is not enough: Understanding the relation between language and spatial cognition. *Child Development*.

Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science*, *14*, 1417–1430.

Shusterman, A., Lee, S., & Spelke, E. S. (2011). Cognitive effects of language on human navigation. *Cognition*, *120*, 186–201.

Simms, N., & Gentner, D. (2008). Spatial language and landmark use: Can 3-, 4-, and 5-year-olds find the middle? In 29th Annual Conference of the Cognitive Science Society (pp. 191–196). Washington, D.C.: Lawrence Erlbaum Associates.

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development*, *85*, 1062–1076.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250.

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