Published in final edited form as:
J Exp Child Psychol. 2018 August ; 172: 107–129. doi:10.1016/j.jecp.2018.02.006.

Children's attention to task-relevant information accounts for relations between language and spatial cognition

Hilary E. Miller and **Vanessa R. Simmering**
Department of Psychology, McPherson Eye Research Institute, Waisman Center, University of Wisconsin-Madison

Spatial skills are foundational for higher cognition and encompass many behaviors, including perceiving and remembering locations, reasoning about object relations, maintaining orientation and navigating through the environment. Early measures of spatial processing predict later math performance (Lauer & Lourenco, 2016; Verdine et al., 2014), and spatial skills in high school predict entry into Science, Technology, Engineering, and Math (STEM) fields (e.g., Wai, Lubinski, & Benbow, 2009). Individual and sex differences in spatial cognition arise early in development (e.g., Linn & Petersen, 1985; Pruden, Levine, & Huttenlocher, 2011) leading to cascading individual differences in prerequisite skills necessary for later life achievement. Despite the importance of spatial skills, there is limited understanding of the mechanisms underlying spatial development, inhibiting our abilities to promote spatial skills.

Multiple studies have found language effects on different spatial skills during early childhood (i.e., 3 to 6 years of age), including searching after disorientation (Hermer-Vazquez, Moffet, & Munkholm, 2001), forming analogies between object relations (Loewenstein & Gentner, 2005), remembering relations among object parts (Dessalegn & Landau, 2008, 2013), mental rotation (Pruden et al., 2011) and reference frame selection in recall (Miller, Patterson, & Simmering, 2016). In interpreting these effects, theorists often propose that language development has a causal impact on spatial cognition through verbal encoding. Some theorists propose that language is essential for spatial development, fundamentally changing our spatial processing (e.g., Shusterman & Spelke, 2005). More commonly, other theorists argue that language facilitates spatial cognition because it directs attention and enhances encoding of relevant spatial information (e.g., Pruden et al., 2011). In this paper, we present an alternative perspective suggesting that verbal encoding does not play a central role in spatial cognition. Rather, children's basic attention skills is a central factor supporting both their language use and spatial skills.

Corresponding Author: Hilary E. Miller, Department of Psychology, University of Wisconsin – Madison, 1202 W. Johnson St., Madison, WI 53706, hemiller@wisc.edu.

Subsets of the data were published in the *Proceedings of the 38th Annual Conference of the Cognitive Science Society* and were presented at the 38th Annual Conference of the Cognitive Science Society, the 72nd Biennial Meeting of the Society for Research in Child Development, and the 9th Biennial Meeting of the Cognitive Development Society.

Most research on language and spatial skills is consistent with multiple explanations, as the studies were not designed to differentiate these alternatives. Prior investigations have typically been correlational, assessing whether children's production of specific terms predicts their spatial performance, or experimental, providing children with specific words to enhance their spatial performance. In both study types, the language measures or manipulations could relate to children's selective attention to relevant information within the tasks, rather than specifically relating to their language use. Selective attention is the ability to filter irrelevant information and focus on relevant information. Selective attention improves during early childhood (e.g., DeMarie-Dreblow & Miller, 1988), the same time that spatial skills are improving (Vasilyeva & Lourenco, 2010) and the age range when many of the links to language have been shown. There has been limited direct investigation of selective attention in relation to effects of language on spatial skills, making it difficult to evaluate possible underlying causal relations. The current research takes a step in this direction by examining individual differences in both language use and selective attention as they relate to young children's spatial performance.

Investigations of language effects on spatial skills

One way that researchers have shown effect of language is through testing children's production of spatial words and how that predicts spatial performance. Children's spatial word production (e.g., "left/right", "by/next", "middle") predicts performance on tasks involving those spatial relations (Hermer-Vazquez et al., 2001; Miller et al., 2016; Simms & Gentner, 2008). Similarly, children who hear more spatial words from their caregivers in unstructured play (Pruden et al., 2011) or in museum-based interactions (Polinsky, Perez, Grehl, & McCrink, 2017) tend to produce more spatial words themselves, and perform better on spatial tasks. Based on these findings, theorists proposed that, as children acquire or gain experience using spatial words, they become better at using language to encode relevant spatial features, enhancing their spatial performance (e.g., Pruden et al., 2011).

Researchers have also tested the effects of experimenter-provided language on children's spatial performance. Children perform better when they hear task-relevant spatial words (e.g., "the red is on the left") before or during spatial tasks, relative to control conditions (e.g., Dessalegn & Landau, 2008; Loewenstein & Gentner, 2005; Miller et al., 2016). These results, along with the correlational results described above, have been interpreted as evidence for language playing a causal role in spatial development either through changing spatial representations or directing attention/encoding (Dessalegn & Landau, 2008; Hermer-Vazquez et al., 2001).

However, other studies showed that language provided to children does not need to be spatial to support their performance (Dessalegn & Landau, 2013; Shusterman, Lee, & Spelke, 2011). Rather, language is useful if it draws attention to information that is relevant for the task. For example, Dessalegn and Landau (2013) tested 4-year-olds in a feature binding task requiring memory for spatial relations between two colored halves of a square (cf. Figure 4c below). When tested with a 1 s delay, children who heard task-relevant language—highlighting the asymmetric relation (e.g., "the black one is prettier")—performed better than children in a control group. A prior study with this task showed no such benefit for a verbal label (e.g., “this is a dax”) or spatial term (e.g., “the red is touching the black”) that did not draw attention to the asymmetry (Dessalegn & Landau, 2008). These results suggest that spatial language is not uniquely helpful; rather any task-relevant language (spatial and non-spatial) that directs attention can support spatial skills.

Limitations of hypotheses positing central role of language

Although the studies reviewed are consistent with hypotheses positing that language facilitates spatial cognition through directing attention, some empirical evidence points to limitations in the extent to which such hypotheses can explain children’s spatial performance. According to hypotheses reviewed, once children can produce task-relevant language (whether spatial or not), they should attend to relevant relations and perform better on spatial tasks (e.g., Pruden et al., 2011). However, young children may know task-relevant words (as in the “prettier” example above) but not use them within the task. Farran and O’Leary (2016) tested children in the feature binding task (described above) first in a baseline condition (no language provided), then with language provided by an experimenter. Only children with knowledge of particular spatial terms (e.g., “left” and “right”) improved when an experimenter provided the terms. Thus, children who knew the relevant spatial terms did not spontaneously use language to support their performance, possibly because they were not attending to that information. Additionally, Plumert and Nichols-Whitehead (2007) tested children’s preferences for using color, size, and location information to disambiguate object locations in a dollhouse. On trials when only location disambiguated the alternatives, 4-year-olds often did not produce location terms spontaneously, but did when prompted by an experimenter. This shows that children could produce location information but likely were not spontaneously identifying that information as relevant prior to prompting. Together, these results suggest that knowledge of language alone is insufficient to account for children’s spatial performance.

Miller, Vlach, and Simmering (2017) more directly tested this hypothesis that language knowledge alone does not account for children’s spatial skills. They argued that, in addition to producing task-relevant words, children must use their language knowledge in task-relevant ways to support their performance. To test this account, Miller et al. (2017) assessed whether children’s task-relevant language use was more predictive of their spatial performance than the number of potentially task-relevant words they produced. Four-year-olds completed a spatial scene description task similar to the task from Plumert and Nichols-Whitehead (2007). The description task was designed to compare the quantity of potentially relevant spatial and non-spatial words produced versus use of these words in task-relevant ways. In the task, children were asked to describe the location of a mouse (see Figure 2 below). Three types of cues (color, size, or location) could describe the mouse’s location, and the relevance of color and size cues to disambiguate the location varied across scenes (cf. Figure 2a and Figure 2d; described further below).

Miller et al. (2017) found that children who produced more relevant than irrelevant cues (both spatial and non-spatial) performed better on a set of spatial tasks. This effect held even when controlling for factors previously shown to predict spatial performance: age, gender, vocabulary (both general receptive and spatial productive), and importantly the quantity of task-related spatial and non-spatial words produced. This result is similar to studies providing children with task-relevant non-spatial cues (e.g., Dessalegn & Landau, 2013), but demonstrates that children's own identification of which cues are most relevant predicts their spatial performance. Miller et al. concluded that the relevance of children's language production was central to predicting their spatial skills. They also speculated that this effect found with task-relevant language use might reflect children's attention to relevant dimensions across tasks, and not specifically abilities to use language.

Building from Miller et al. (2017), we propose an alternative account explaining relations between language use and spatial cognition. Children need to know when particular words are relevant for the words to be used within a task (Miller et al., 2017). Thus there are two essential elements for children to use language in task-relevant ways: they need to produce the task-related words (words potentially relevant for the task) and selectively attend to task-relevant information. Accordingly, the element of children's task-relevant language use that is most related to their spatial performance is selective attention, not their abilities to produce the task-related words. That is, selective attention on its own predicts spatial performance. Children's use of task-relevant language is only associated with spatial performance because of this shared reliance on selective attention (Figure 1a). Thus, correlations between children's language use and spatial skills does not mean that children are necessarily using verbal encoding while performing spatial tasks. This is consistent with evidence that children can perform complex spatial tasks before they have words for verbal encoding (for review, Newcombe & Ratliff, 2007), and that young children do not reliably use verbal strategies in memory tasks until 8 years of age (Pickering, 2001).

Because past research has not included non-verbal measures of attention, it remains an open question whether effects of task-relevant language use on spatial performance occur because of shared reliance on selective attention or because language is playing a causal role in spatial performance. Considering that children need to attend to relevant information to use language in relevant ways, it is possible that selective attention is causing children to more effectively use language within the spatial task. This use of language could improve children's encoding of spatial information, in turn enhancing their spatial performance (Figure 1b). This would be consistent with a verbal encoding hypothesis, with the nuanced element that selective attention is facilitating children's language use. When children are provided with language cues by an adult, the language can direct their attention to improve their spatial performance (e.g., Dessalegn & Landau, 2008). However when children are on their own, they need to attend to the relevant information to use language in relevant ways as children do not spontaneously use their language knowledge to enhance their spatial performance (Farran & O'Leary, 2016; Miller et al., 2017).

Current Studies

The current experiments were designed to understand the role that selective attention plays in children's spatial performance, motivated by our hypothesis that attention supports both spatial skills and task-relevant language use. We extend the literature by measuring children's selective attention as it relates to both their language use and their spatial performance. Experiment 1 follows the design of Miller et al. (2017) with the addition of a selective attention task. This experiment tested whether 4-year-old children's adaptive language use relates to their spatial performance primarily because these abilities are supported by selective attention (Figure 1a), or language is uniquely influencing spatial skills as predicted by a verbal encoding hypothesis (Figure 1b). Experiment 2 tested whether performance in the description and attention tasks was specific to the designs from Experiment 1 by testing variations of these tasks. We focus on 4-year-olds because this is an age range where robust effects of language and spatial cognition have been found (e.g., Dessalegn & Landau, 2008; Miller et al., 2017; Shusterman et al., 2011). Additionally, this is an age range where there is substantial variance in children's performance on the spatial and language measures used in this investigation (Miller et al., 2017).

Experiment 1

Experiment 1 tested whether selective attention relates to associations between children's adaptive language use and spatial performance. Our design follows Miller et al. (2017) using their spatial scene description task (description task) to assess adaptive language use and task-related language production. We used the spatial tasks (Children's Mental Transformation Task, Spatial Analogies, and Feature Binding) previously found to be most associated with children's adaptive language use. These tasks do not rely on a single underlying spatial ability, but rather tap multiple spatial skills and require attending to different dimensions across trials and tasks. As a non-verbal measure of selective attention, we added a spatial scene attention task (attention task) assessing children's attention to the cues from the description task (see Figure 3). Children were presented with the same spatial scenes from the description task, and then identified the color, size, or location of the referent object the mouse was on after a brief delay. This task taps selective attention as children needed to selectively attend to properties specific to the object the mouse was on in a scene with other referent objects. Although this task involves memory due to the delay, we believe it provides an index of attention because memory demands are minimal: young children demonstrate robust recognition of object features following short delays, even with multiple features per object (Simmering & Wood, 2017). Additionally, children's selective attention has been assessed using memory tasks (e.g., DeMarie-Dreblow & Miller, 1988) and children's selective attention at encoding relates to their later recognition (Blumberg, Torenberg, & Randall, 2005; Roderer, Krebs, Schmid, & Roebers, 2012).

Experiment 1 had two goals. The first goal was to understand the extent to which performance in the description task from Miller et al. (2017) related to children's selective attention. We evaluated this in two ways. First, we investigated whether children attended to the cue types similarly across the description and attention tasks. We predicted similar patterns across tasks in which cues children were likely to produce (in the description task) and recognize (in the attention task). Second, we assessed whether individual differences in children's adaptive language use was predicted by both their production of task-related language and their attention score. We hypothesized that children's adaptive language use would be predicted by both their production of potentially task-related language and their selective attention skills (see Figure 1 white arrows). Although these analyses do not differentiate the two hypotheses presented in Figure 1, they were motivated by our hypothesis that children's task-relevant language production involves components of selective attention.

The second and central goal was to test whether selective attention relates to associations between adaptive language use and spatial performance. Using hierarchical regression, we first tested whether children's adaptive language use predicted their spatial performance beyond their age, gender, and general receptive vocabulary, which it did. We then added selective attention to the regression model in a second step to test whether it accounted for the variance shared between adaptive language use and spatial performance. This analysis of shared variance was designed to test whether our results are most consistent with the hypothesis that attention supports both spatial skills and task-relevant language use (attention accounts for variance previously associated with adaptive language use) or with a hypothesis positing a unique role of language (adaptive language accounts for unique variance).

Method

Participants—Sixty-eight 4-year-old children (*M* age=4.51 years, *SD*=0.40, 31 girls) participated. An additional 14 children participated but were excluded due to: incomplete data (4), experimenter error (1), technical problems (2), parental interference (1), non-compliance (1), not talking during the description task (3), insufficient vocabulary (1, described below), or as an outlier in the regression model (1, described below). Participants were recruited from a database compiled by a university research center and were primarily from white middle-class backgrounds (individual demographic information was not collected) and received prizes for participation.

Design and Procedures—Children were tested individually in the description task, attention task, three spatial tasks (listed below), 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 spatial tasks and the second and fourth tasks were the description and attention tasks (counter-balanced across participants), and the final task was the PPVT-IV. Children received small prizes and were offered breaks between tasks. Caregivers completed a productive spatial vocabulary checklist. Sessions lasted approximately 45 minutes.

Spatial Scene Description Task: The same procedures were used as reported in Miller et al. (2017). Before the description task, children completed a warm-up task to help them understand the task and feel comfortable talking aloud. In the warm-up task, children viewed 10 PowerPoint slides showing one familiar object per slide and were asked to describe what they saw to a stuffed animal who was not facing the screen ("Tell Bucky what you see"). The experimenter discouraged pointing during the warm-up and description tasks by instructing children to sit on their hands and use their words.

The description task tested how children verbally disambiguated a target referent object relative to other objects (see Figure 2). The task was conducted in PowerPoint. Each trial showed a picture of a spatial scene centered on the slide. Each scene included three referent objects (e.g., beds) distributed diagonally across the screen; the diagonal orientation (top-left to bottom-right vs. top-right to bottom-left) varied randomly across trials. This alignment allowed children to use a variety of spatial words to describe the mouse's location (e.g., front, middle, back, first, last, left, right, top, center, bottom). The mouse was in a support relation to the target referent object. On each trial, children were asked to "Tell Bucky where the mouse is" to a stuffed animal not facing the screen. Across trials, the scenes varied in the number of relevant cues for describing the target referent object: 3-cue (color, size, and location, Figure 2a), 2-cue-color (color and location, Figure 2b), 2-cue-size (size and location, Figures 2c), or 1-cue (location, Figure 2d). The description task included 24 trials, 6 per scene type with 2 trials for each mouse's location (i.e., front, middle, back), presented in one of four pre-determined randomized orders.

Spatial Scene Attention Task: The attention task was modeled after the description task and tested children's attention to properties of the target referent object by probing recognition after a short delay (see Figure 3). The scenes were identical to those in the description task and presented in the same order. Following presentation of the scene, there was a 1 second delay with a blank screen, and then a test array was presented. Test arrays probed recognition for one dimension (color, size, or location) of the target referent object. 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 the mouse was on ("which bed was the mouse on?"). 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 3a). The structure of this task was modeled after the feature binding task (Dessalegn & Landau, 2008), which also used a 1 second delay and test arrays with three options. For each scene type (e.g., 2-cue-color), children were presented twice with each probe type. The probe dimensions were randomized across trials and children were unaware of which dimension would be probed.

Children completed three practice trials before the 24 test trials. During the practice trials, if the child missed a trial, the experimenter would re-present the scene and say "look the mouse was not on this one, let's try again," repeating the trial until the child responded correctly. The experimenter did not use any color, size, or location terms to avoid encouraging a verbal strategy. On the test trials, no accuracy feedback was provided.

Spatial Tasks: Children participated in short versions of three spatial tasks (see Figure 4): Children's Mental Transformation Task (CMTT, Huttenlocher & Levine, 1990; Pruden et al., 2011), Spatial Analogies Task (Levine, Huttenlocher, Taylor, & Langrock, 1999; Pruden et al., 2011), and Feature Binding Task (Dessalegn & Landau, 2008). These tasks were chosen because they have previously been shown to relate to children's spatial language, and were included in Miller et al. (2017) for comparison with children's adaptive language use. Furthermore, they require children to attend to different dimensions (e.g., different orientations, different relations) across trials and therefore do not rely on one specific spatial ability. For the CMTT and Spatial Analogies Task, we used the short version from Pruden et al. (2011) and for the Feature Binding Task, we used half the number of trials as in Dessalegn and Landau (2008). In the CMTT (10 trials; Figure 4a), children saw two pieces of a shape and selected which of four shapes the two pieces make if combined. In the Spatial

Analogies Task (13 trials; Figure 4b), children viewed a picture depicting two objects in a spatial relation and chose which of four pictures shared that relation. In the Feature Binding Task (12 trials; Figure 4c), children saw a square with different colors on opposite sides and chose a matching figure out of three options following a 1 second delay.

Vocabulary assessments: The PPVT-IV measures receptive vocabulary and involved children pointing to one of four pictures depicting the target word. The spatial vocabulary checklist measured productive spatial vocabulary and included 80 words (Miller et al., 2017). Caregivers indicated whether they have heard their child produce these words spatially.

Coding and Measurement

Spatial scene description task: Each session was transcribed by two research assistants and coded by two additional research assistants, all naive to the hypotheses. Transcribers agreed on 89% of trials and discrepancies were resolved by a third research assistant blind to the first two transcripts. Coders agreed on 98% of the dimensions coded, and disagreements were resolved by the first author. Coders scored the number of times children mentioned color, size, or location terms (referent's location in the scene, e.g., *top chair*) and whether they used the terms correctly.

For goal 1, we calculated the average number of unique color, size, and location terms correctly used per trial by scene type during the description task for the task comparison analysis. We also averaged the unique usages of these terms together to create the task-related production score for the regression analysis. For goals 1 and 2, we created an adaptation score by calculating each child's production of relevant versus irrelevant cues during the description task. On each trial, a response was scored for the three cue types (color, size, and location) as a 0 or 1 depending on whether children produced those cues. Multiple terms referring to the same cue type were only counted as 1 (i.e., "front corner" and "front" would both count as 1). Color and size terms were coded as negative (irrelevant) if produced on trials when the terms could not differentiate the referents (color negative on 2-cue-size and 1-cue trials [Figure 2c and 2d], and size negative on 2-cue-color and 1-cue trials [Figure 2b and 2d]). On all other trials color and size cues were coded as positive (relevant), and location cues were always coded as positive because location was always relevant. Adaptation scores could range from −2 to 3, and were positive when children provided more relevant than irrelevant cues and negative when provided more irrelevant than relevant cues. Scores closer to 0 reflected performance that did not differ by type (e.g., mentioning color terms on all trials, not mentioning any relevant cues). Thus, the adaptation score reflected the context in which children provided color and size cues, reflecting their use of task-relevant language across trials.

To ensure that our analyses only included children with sufficient language knowledge, we made two types of exclusions. We excluded trials if the caregiver indicated on the checklist that their child does not produce any of the potentially relevant words for describing the referent object's location (e.g., if "middle" was not checked, we excluded trials when the mouse was on the middle object). However, if children correctly produced a word during the task that the caregiver did not check, we included such trials as the child demonstrated knowledge of the relevant word (e.g., if the child produced “middle” correctly, we included all trials with the mouse in the middle location). 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 at the bottom). These criteria resulted in the exclusion of 124 trials (8% of total trials; 52 trials [3% of total] for incorrect word use and 72 trials [5% of total] for relevant location words not checked by caregiver) from 24 children’s data, with one child’s data being entirely excluded due to the caregiver indicating and the child not producing any of the relevant words.

Spatial scene attention task and spatial tasks: The experimenter marked responses on a session sheet during participation. 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 4). 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. The spatial composite score was created because we were interested in children’s general ability to solve spatial tasks and adapt to demands across tasks, rather than a single underlying skill (e.g., mental rotation, relational reasoning, visualization). For the attention task, we calculated the proportion of correct trials by probe dimension and scene type for the goal 1’s task comparison analyses and the total proportion of correct trials for goals 1 and 2’s regression analyses.

Vocabulary assessments: The experimenter marked 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). Standardized scores were calculated offline using established norms (Dunn & Dunn, 2007).

Results and Discussion

Experiment 1 investigated relations between 4-year-olds’ spatial skills, adaptive language use, and selective attention. One participant was removed from analyses for having high influence on the regression model, based on Cook’s d outlier analysis (Fox, 1991). Tables 1 and 2 present descriptive and correlational statistics of our measures used in the regression analyses. As shown in Table 2, all measures were positively correlated with the spatial composite score, except for gender in which we found no effect. As a preliminary analysis, we tested whether children’s adaptation score, attention score or spatial composite scores differed based on order (i.e., completing the description or attention task first), but found no significant differences ($ps > .347$).

Spatial scene description and attention tasks—The first goal was to understand the extent to which performance in the description task related to children’s selective attention. First, we compared the types of cues that children produced and recognized when identifying object locations in spatial scenes. For the description task, we assessed whether children produced different numbers of color, size, and location terms overall, and whether their production differed based on the cue variability in the spatial scenes (cf. Figure 2). As Figure 5A shows, children produced color terms most and location terms least. We conducted a Friedman’s test to compare the proportion of trials on which children mentioned each cue type (color, size, and location, collapsing across scene types) and found a significant effect, $\chi^2(2, N=68)=49.19$, $p<.001$. Pairwise comparisons showed that children mentioned color ($M=.41$, $SD=.40$) and size ($M=.24$, $SD=.32$) significantly more than location ($M=.05$, $SD=.18$); color and size did not differ significantly ($p=.072$).

To compare production of color, size, and location terms across scene types (3-cue, 2-cue-color, 2-cue-size, and 1-cue), we conducted separate Friedman’s tests for each cue type. There were significant differences across scene types for all cues: color $\chi^2(3, N=68)=45.47$, $p<.001$; size $\chi^2(3, N=68)=43.05$, $p<.001$; location $\chi^2(3, N=68)=12.15$, $p=.007$. Pairwise comparisons following up on these effects showed different patterns across cue types (see Figure 5A for means). Children produced more color terms when color varied in the scene (3-cue and 2-cue color) than when it did not (2-cue size and 1-cue trials). Children produced size terms more when it varied and color did not (2-cue-size trials) than on any other trial type. For location terms, no pairwise comparison reached significance (ps>.319). Overall, these results show that children were sensitive to cue variability in the scenes, as they produced more color and size terms when those terms were relevant, similarly to finding of Plumert and Nichols-Whitehead (2007). However, they rarely produced location terms despite those terms being relevant across all trials.

For the attention task, we analyzed whether children’s performance differed based on the dimension probed and the types of cues in the scene. As Figure 5b shows, children recognized the same types of cues that they tended to produce during the description task, with best performance on color probes and worst on location probes (cf. Figure 5a). We conducted a two-way ANOVA on proportion correct, with probe dimension (color, size, and location) and scene type (3-cue, 2-cue-color, 2-cue-size, and 1-cue) as within-subjects factors. The ANOVA revealed main effects of probe dimension ($F_{2,134}=90.73$, $p<.001$, $\eta_p^2=.575$) and scene type ($F_{3,201}=37.15$, $p<.001$, $\eta_p^2=.375$), which were subsumed by a two-way interaction ($F_{6,402}=13.25$, $p<.001$, $\eta_p^2=.165$). Bonferroni post-hoc comparisons of probe dimensions revealed significant differences in each pair-wise comparison (color $M=.83$, $SD=.18$; size $M=.61$, $SD=.21$; location $M=.42$, $SD=.18$). Bonferroni post-hoc comparisons of scene type revealed that children performed better on 2-cue-size ($M=.73$, $SD=.18$) and 1-cue ($M=.69$, $SD=.17$) than on 3-cue ($M=.53$, $SD=.17$) and 2-cue-color scene types ($M=.54$, $SD=.18$); no other differences were significant.

To understand the interaction, we conducted one-way ANOVAs on proportion correct, with scene type as a within-subject factor, separately for each probed dimension. These analyses revealed no effect for color probes ($p=.267$), but significant differences for size ($F_{3,201}=37.79$, $p<.001$, $\eta_p^2=.361$), and location probes ($F_{3,201}=12.27$, $p<.001$, $\eta_p^2=.155$; see Figure 5b for means). Bonferroni post-hoc comparisons following up on the main effects of size probes revealed higher performance on 2-cue-size and 1-cue scenes than 3-cue and 2-cue-color scenes. Performance was also higher on 3-cue than 2-cue-color scenes; no other differences were significant. Bonferroni post-hoc comparisons following up on the main effect of location indicated lower performance on 3-cue scenes than all others; no other differences were significant. These results show that overall across trials children were best at recognizing color and performed worse on trial types when color varied in the scene than when color did not vary. Specifically, children's recognition of color was not affected by variation in size, or location in the scenes. However for size, children were better at recognizing size only when color did not differ in the scene. Children's overall recognition of location was poor, especially when color and size both varied in the scene.

As hypothesized, children's production and recognition of the cue types were similar across tasks. Color was produced and recognized the most, with production and recognition of size and location terms depending on whether color varied in the scenes. These results demonstrate that children were not always selectively attending to relevant cues in the description and attention tasks, having an inclination to attend to color over the other cue types. The similarities in performance across these tasks suggest that the two tasks at least partially assessed the same latent ability of selective attention as predicted by our hypothesis.

To further examine how selective attention related to children's adaptive language use, we investigated factors predicting the adaptation score. As shown in Figure 1 (white arrows), children's adaptation scores should reflect both their task-related language production and their selective attention. As predicted, both the task-related production (t_{62}=4.73, p<.001, ΔR^2=.171) and the attention score (t_{62}=2.48, p=.018, ΔR^2=.047) uniquely predicted children's adaptation score controlling for age, gender, and PPVT-IV score (see Table 3). This result shows that basic attentional mechanisms play a role in children's use of language in task-relevant ways.

Predictors of spatial skills—Our second goal was to evaluate how selective attention relates to spatial skills and relations between adaptive language use and spatial performance. We used regression analyses to first test whether children's adaptive language use predicted their spatial performance above and beyond age, gender, and PPVT-IV score. This analysis was conducted to confirm that effects of the adaptation score predicting spatial performance held in this sample (Miller et al., 2017). As predicted, we found that children's adaptation scores significantly predicted their spatial composite scores (t_{63}=2.36, p=.021, ΔR^2=.049) when controlling for age, gender, and PPVT-IV score[1]. We next added the attention score to the model to assess whether selective attention accounts for the variance in spatial performance previously accounted for by adaptive language use. We found that selective attention significantly predicted children's spatial composite score (t_{62}=2.18, p=.033, ΔR^2=.040) as shown in Figure 6 and Table 4, and the variance accounted for by the adaptation score was no longer significant (p=.118).[2,3] A follow-up analysis of this effect of the attention measure showed that the spatial composite score was most predicted by the

---

[1] This analysis is not an exact replication of Miller et al. (2017) due to differing research questions. Miller et al. asked whether children's adaptive language use predicted spatial performance beyond their spatial word production, leading them to control for both children's spatial vocabulary (parent checklist) and quantity of task-related language production (both non-spatial and spatial). Using the same analysis, our results replicate theirs, t_{60}=2.21, p=.031, ΔR^2=.045. However, the current paper is concerned with the extent to which the effect of the adaptation score reflects children's language use and thus we do not control for children's spatial vocabulary and task-related language production to avoid reducing variance in the adaptation score that reflected children's language abilities.

attention task trials in which color and size were differentiated in the scene and children were probed on the differentiated dimension (i.e., color probe trials for 3-cue and 2-cue-color scenes and size probe trials for the 3-cue and 2-cue-size scenes; see Miller & Simmering, 2016, for complete details). These trials likely reflect selective attention more than trials in which the probed dimension was not differentiated in the scene as children specifically needed to attend to the target referent object to encode the cues. In undifferentiated trials, it is possible to encode the color or size from all of the referent objects. These results are consistent with our hypothesis showing that children's selective attention is related to their spatial performance, accounting for variance previously thought to relate specifically to children's language use.

It is important to recognize that mediation analyses are often conducted when the effect of one variable is decreased by the addition of another to a regression model (MacKinnon, Krull, & Lockwood, 2000), similar to our results above. Mediation analyses can add to hierarchical regression because they reveal insights into the causal direction between a predictor and outcome variable. In a mediated relation, the predictor variable causes the mediator, which in turn causes the outcome variable. As an illustration in this domain, Pruden and Levine (2017) found that parents' spatial word use when children were 14–26 months mediated the relation between children's sex and spatial word usage at 34–46 months. In their model the effect of sex on children's spatial word usage was reduced when parents' spatial word usage was added to the model.

In our model, this could mean that the attention score mediated the relation between the adaptation score and the spatial composite score, since its addition reduced the variance accounted for by the adaptation score. However, mediation would imply that the adaptation score causes selective attention, analogous to how children's sex leads parents to differentially use spatial words in Pruden and Levine's (2017) study. However, both conceptually and as found in the regression analysis from goal 1 (Table 3), adaptive language use in the description task depends on selectively attending to the task-relevant information. If a causal relation existed, the adaptation score should be the mediator, as selective attention could cause adaptive language (see Figure 1b black arrows). Our data do not support this possible mediation model because the adaptation score does not account for any variance in spatial performance beyond the attention score. The pattern in our regression results is more consistent with a confounding relation (MacKinnon et al., 2000), in which a third variable (attention) accounts for the relation between the predictor (adaptation score) and outcome (spatial composite) variables. Unlike mediation, the third variable is not an intermediate variable and the model does not assume a causal direction among the variables. We contend that our hypothesis that attention supports both spatial skills and task-relevant language use, as well as our regression results, are most consistent with a confounding model.

---

[2] The qualitative pattern of results remained the same with the outlier data included. The only quantitative difference was that the adaptation score was marginal in the first step (p=.054)

[3] We also conducted the same regression analysis separately for each spatial task in two steps, but none of these models reached significance (step 1 adaptation score: ps>.163 and step 2: selective attention score: ps> .071). Due to the small number of trials in each spatial task, there was likely not enough variance to detect such effects.

Overall, the results of Experiment 1 are consistent with our predictions demonstrating that selective attention is associated with both children's spatial performance and adaptive language use. The results suggest that relations found between children's language use and spatial cognition likely result from a third variable of selective attention. We acknowledge that our results are correlational and therefore cannot provide evidence of causation. However, we believe the variance associated with the attention tasks beyond the language effect is most consistent with our hypothesis that attention supports both spatial skills and task-relevant language use and indicates the need for theorists to more thoroughly evaluate the processes that could underlie correlations across tasks and results from experimental manipulations.

This need to consider how tasks reflect underlying processes raises an important question about our results: could children's performance in the attention and description tasks be driven by demands specific to our task designs? We interpreted results from these tasks as reflecting individual differences in selective attention, but it is important to evaluate whether the same performance patterns emerge when task demands vary. In particular, there are three questions we sought to address in an additional experiment. First, were children influenced to perform these tasks in similar ways through our within-subjects design? We found no systematic differences based on order, but including both tasks might have led children to approach them more similarly than if tested alone.

Second, did the unpredictability of the probe dimension in the attention task lead children to adopt task-specific strategies? Because the probed dimensions varied randomly, the optimal strategy was to remember the color, size, and location of the target referent object on every trial. Although prior studies suggest that children can remember two features of the same object with little or no additional cost depending on how memory is probed (Riggs, Simpson, & Potts, 2011; Simmering & Wood, 2017), those studies specifically instructed children to attend to shape and color, and did not probe location. It is possible that children found it difficult to attend to all dimensions sufficiently to support recognition in our task, leading them to prioritize only one dimension instead of attending to multiple dimensions.

Third, could children's descriptions be limited by the need to produce certain types of words? Production is more demanding than comprehension (Clark & Hecht, 1983), which can make it difficult to interpret language production as indexing underlying language abilities. In our description task, it is possible performance patterns reflected the difficulty of retrieving relevant words, rather than indexing whether children attended to relevant dimensions. We partly controlled for this by not including trials on which the participant's caregiver indicated they did not produce the relevant words. However, children may have struggled in retrieving the relevant words in our task even if they could produce them in other contexts. In Experiment 2, we followed up on these questions to further understand the factors influencing performance.

Experiment 2

Experiment 2 addressed the questions raised from Experiment 1 by testing whether 4-year-olds' performance in the attention and description tasks reflected task-specific demands or selective attention. To avoid interactions between tasks, we tested separate groups of children in the attention task (Experiment 2a) and the description task (Experiment 2b). Experiment 2a tested performance on different probe dimensions in the attention task across between-subjects conditions, such that each child was probed along a single dimension. This allowed us to test whether a predictable probe dimension would systematically bias children to attend to that dimension. Experiment 2b tested children in the description task with reduced retrieval demands for production by providing children with cards representing cue dimensions and asking them to select the cue(s).

Our prediction for Experiment 2a was that children would show the same general pattern as in Experiment 1, reflecting the influence of attention on recognition of the cues across scenes. Alternatively, if performance in Experiment 1 resulted from the unpredictability of the probe dimension, then children's performance in Experiment 2a should be higher than in Experiment 1. Our prediction for Experiment 2b was that children would select similar cues as produced in the description task from Experiment 1 (and found in Miller et al., 2017), which were not always the most relevant cues. This would reflect children's selective attention to the cues, not just their ability to produce particular words. This hypothesis is also consistent with prior research showing that young children often fail to notice ambiguity in both their own (Plumert & Nichols-Whitehead, 2007) and others' descriptions (Plumert, 1996). However, if prior performance in the description task primarily reflected demands on retrieving relevant words, then children should perform better in Experiment 2b where production demands were eliminated.

Experiment 2a Method

Participants—Sixty-two 4-year-old children (*M* age=4.47 years, *SD*=.45, 25 girls) participated. An additional 8 children participated but were excluded due to: not understanding the task (6) and non-compliance (2). Participants were recruited through community preschools and were primarily from white middle-class backgrounds (individual demographic information not collected). Individual children were not compensated, but preschools received gifts as a thank-you for participating.

Design and Procedures—Children were tested individually in a quiet space in their preschool. Children participated in a modified version of the attention task from Experiment 1 in which only one dimension of the target referent object (color, size, or location) was probed across all 24 trials. The scenes were identical across all conditions and were presented in one of two randomized trial orders per probe condition. Prior to participation, children were randomly assigned to one probe dimension condition, with the final sample including 21 children in color-probe, 22 children in size-probe, and 19 children in location-probe conditions. The same procedures for the practice and test trials were used as in the attention task from Experiment 1.

Experiment 2a Results and Discussion

The goal of Experiment 2a was to evaluate whether performance on the attention task in Experiment 1 reflected individual variations in how children attended to task-relevant information or reflected demands specific to the experimental design of the probe dimension being unpredictable. As Figure 7 shows, children's performance in Experiment 2a closely aligned with that of Experiment 1. We conducted similar analyses to Experiment 1 by testing whether children's performance varied based on both probe dimension and scene type. We conducted a two-way mixed effects ANOVA on proportion correct, with probe dimension (color, size, and location) as a between-subjects factor and scene type (3-cue, 2-cue-color, 2-cue-size, and 1-cue) as a within-subjects factor. The ANOVA revealed main effects of probe dimension ($F_{2,59}$=32.52, p<.001, $\eta_p{}^2$=.524) and scene type ($F_{3,177}$=16.12, p<.001, $\eta_p{}^2$=.215); the interaction did not reach significance (p=.060). Bonferroni post-hoc comparisons for the probe dimension effect revealed significant differences between all conditions (color M=.86, SD=.10; size M=.69, SD=.16; location M=.46, SD=.20). Bonferroni post-hoc comparisons for the scene type effect revealed that children performed better on 2-cue-size (M=.76, SD=.27) and 1-cue (M=.72, SD=.25) than on 3-cue (M=.65, SD=.26) and 2-cue-color (M=.59, SD=.26) scene types.

As predicted, the results of Experiment 2a parallel those of the attention task of Experiment 1, with the same patterns across probed dimensions and generally lower performance when color varied in the scene. Although we did not compare the results statistically between experiments (due to the difference in within-versus between-subject designs), the similar patterns suggest that children's performance in Experiment 1 was not strongly influenced by including the description task or the unpredictable probe dimension. Even when size or location were probed on every trial, children in the current experiment performed worse than those probed on color. Furthermore, differentiation of color in the scenes still influenced performance for children who were never probed on that dimension.

Experiment 2b Method

Participants—Thirty 4-year-old (M age=4.41 years, SD=.41, 16 girls) children participated. One additional child participated but was excluded for being non-compliant. Participant recruitment and preschool compensation were the same as in Experiment 2a.

Design and Procedures—Children were tested individually in a quiet space in their preschool. Children participated in a cue-selection task assessing whether they adaptively selected relevant cues for describing locations in a spatial scene. This task was modeled after the description task in Experiment 1 and Miller et al. (2017) but did not require children to produce words. Instead, children were presented with three cards that corresponded to the potentially relevant cues and were asked to select the cards they would use to describe the target referent objects.

Children were first presented with 9 cards indicating the range of colors, sizes, and locations used in the study (see Figure 8a). Each color card had a square in the middle that showed red, blue, or yellow. Each size card had a stick-figure person in the middle that was small, medium, or large. Each location card had a house on it that was located in the bottom left, middle, or top right. To familiarize the children with the cards, the experimenter first presented the color cards and labeled them. Then the experimenter randomly sorted the three cards in a different order and asked the child to give each card to a stuffed animal that was not looking at the cards (e.g., "if you want to give the frog the blue card which one would you give.”). The same procedure was repeated for the size and location cards. Then the experimenter laid out three cards of different dimensions (e.g., red card, big card, middle card) and asked the child to give the stuffed animal each card and repeated the procedure until each of the 9 cues had been chosen. Finally, the experimenter presented the child with three cards of different dimensions and asked the child to retrieve 2 of the dimensions and then 3 of the dimensions to indicate to the child that they could choose more than one dimension (e.g., “give the frog the red card and the big card”). If the child picked the incorrect card during this familiarization task, the experimenter would correct the child and later ask for that card again to ensure that the child understood what the card represented.

After the familiarization task, the experimenter started the cue-selection task. The child was presented with the same spatial scenes used in Experiment 1. Each spatial scene was presented on a separate piece of paper within an easel binder facing the child. The experimenter explained the task as a game in which the child will help the stuffed animal find the mouse. On each trial, the experimenter presented the child with three cards that represented the dimensions of the target referent object the mouse was on (see Figure 8b). The cards always corresponded to the cues of the target referent object. For example, if the mouse was on a small red boot located at the bottom of the page, the experimenter would present the child with the red, small, and bottom cards (Figure 8b). The order of card presentation was pseudo-randomized such that the experimenter presented each dimension card first, second, and third, each on a third of the trials (i.e., the color card was not always presented first). On each trial, the experimenter said to the child “To help the frog find the mouse, would you tell the frog the mouse is on the red boot (pointing to the red card), the bottom boot (pointing to bottom card), the small boot (pointing to the small card)? Remember you can give the mouse one card or you can give the mouse more than one card.” Children could pick up the cards and/or verbally describe their choice(s). The experimenter marked on a session sheet which cue(s) the child picked. During familiarization and throughout the task, the experimenter ensured that the child knew they could give more than one card.

Experiment 2b Results and Discussion

Experiment 2b tested whether children’s descriptions in the cue-selection task, when retrieval demands for producing potentially-relevant cues were eliminated, would be similar to those of the description task from Experiment 1. We first calculated the proportion of trials on which children selected each cue type, similarly to Experiment 1. One notable difference from Experiment 1, however, is that the cue-selection task required children to select at least one cue, whereas children could have provided no cues in Experiment 1 (e.g., saying only “the mouse is on the boot”). As Figure 9 shows, children showed similar performance patterns to Experiment 1, providing mostly color cues and rarely using location cues, but with overall higher levels for all cues due to the requirement of selecting at least one cue.

We first conducted a Friedman’s test to compare the proportion of trials on which children selected the various cue types (color, size, and location) collapsing across trial types. This test showed that children significantly differed in the types of terms selected across trials,

$\chi^2(2, N=30)=24.23$, $p<.001$, with pairwise comparisons showing that children's selection of each cue type differed significantly (color $M=.75$, $SD=.21$; size $M=.51$, $SD=.27$; location $M=.35$, $SD=.29$). These results parallel that of the description task from Experiment 1, as well as the attention task in Experiments 1 and 2a.

We also assessed whether the relevance of cues in the spatial scene influenced children's selections to gain insight into how adaptive children were across trials. We conducted separate Friedman's tests for each cue type to evaluate differences in the rate at which cues were selected across scene types. These analyses showed significant differences across scene types for all cues: color cues $\chi^2(3, N=30)=20.15$, $p<.001$; size cues $\chi^2(3, N=30)=23.80$, $p<.001$; location cues $\chi^2(3, N=30)=11.94$, $p=.008$. Pairwise comparisons following up on each of these effects showed different patterns across cue types (see Figure 9 for means). Children selected color cues more on the 3-cue and 2-cue-color than on the 2-cue-size scene types, and more on the 2-cue-color than cue scene types. For size cues, children selected size more on the 2-cue-size than 2-cue-color and 1-cue scene types, and more on the 3-cue than on the 2-cue-color scene. For location cues, children selected location more on 1-cue than on 2-cue-size scene types, with no other significant differences. These results show that children were somewhat sensitive to the cues in the spatial scene, with their selection of color and size relating to whether those dimensions differed, but children rarely provided location cues despite their relevance.

Overall as predicted, the results from Experiment 2b parallel Experiment 1 with children being somewhat adaptive in providing color and size cues. However, children still overwhelmingly provided color cues despite these cues being relevant on only 50% of trials, and infrequently selected location cues despite their relevance on all trials. We conducted one further analysis to more closely approximate the adaptation scores calculated in Experiment 1. As noted above, one difference between the cue-selection task and the description task was that the cue-selection task but not the description task required children to provide at least one cue by selecting at least one card. This makes it difficult to compare results directly between experiments. To account for this difference, we re-calculated adaptation scores from Experiment 1 after removing trials on which children did not provide color, size, or location cues (722 trials [46%] from 61 participants; note that these were roughly evenly distributed across scene types). This resulted in a mean adaptation score of .44 ($SD=.42$; cf. $M=.29$, $SD=.39$, with these trials included). To ensure that the removal of participants did not significantly alter the relation between adaption scores and children's spatial composite scores, we re-analyzed our data with the adjusted adaption score and found it correlated significantly with children's spatial composite scores ($r_{42}=.40$, $p=.007$). We then calculated the adaptation score for Experiment 2b as in Experiment 1, which resulted in a mean of .55 ($SD=.33$). The difference in task designs precludes statistical comparison of these two scores, but the similar results are consistent with our expectation that production demands were not the primary driving factor behind the results in Experiment 1. Children were only slightly better in Experiment 2b when the potentially relevant cues were provided. These results provide converging evidence in line with our hypothesis that a more general mechanism of selective attentions supports producing language in relevant ways.

General Discussion

The current studies examined the role of selective attention in young children's spatial skills and relations found between language and spatial cognition. Prior research focused on verbal encoding as a mechanism underlying spatial development (e.g., Hermer-Vazquez et al., 2001). However, some research has challenged this perspective (e.g., Miller et al., 2017; Newcombe & Ratliff, 2007) suggesting that non-verbal mechanisms could underlie relations with language. We present a novel hypothesis (Figure 1a), to explain language effects, theorizing that language and spatial cognition are related due to their shared reliance on selective attention.

Experiment 1 examined individual differences in children's selective attention as it relates to their task-relevant language use and spatial performance. Results showed that children produced and recognized similar types of cues across description and attention tasks, suggesting that both tasks assessed similar underlying skills. In terms of factors involved in children's adaptive language use, both the quantity of children's task-related language use (color, size, and location terms) and their selective attention scores uniquely predicted their spatial performance, demonstrating contributions of selective attention in abilities to use task-relevant language. Additionally, as predicted we found that selective attention predicted children's spatial performance beyond effects of adaptive language use as well as age, gender, and general receptive vocabulary. Experiment 2 followed up on Experiment 1 to investigate whether performance in the attention (Experiment 2a) and description tasks (Experiment 2b) reflected task-specific demands or more general demands of selective attention. We found similar performance patterns as in Experiment 1, even after reduced demands of being probed along multiple dimensions (Experiment 2a) or retrieving relevant words (Experiment 2b). These results suggest that children's performance reflected their selective attention across production and attention tasks. Overall, the results of this investigation are consistent with predictions that attention supports both spatial skills and task-relevant language use. This work highlights the importance of considering the role of selective attention in underlying spatial development and accounting for effects previously attributed to children's language use.

Additionally, the current investigation adds to research suggesting that identifying task-relevant cues are important for spatial performance, even if the cues are non-spatial (e.g., Dessalegn & Landau, 2013; Miller et al., 2017). We demonstrated that this type of relevance extends beyond language to more basic attention processes. In the current studies, this type of attention was mostly related to differentiation of color and size cues within a spatial scene. However, this effect is likely not specific to these types of cues (e.g., color) but instead related to attention to task-relevant cues more generally. In the future, it will be essential to test selective attention across other tasks and measures to gain a broader picture of how it relates to language and spatial skills.

Although our results are consistent with the hypothesis that improvements in selective attention support spatial development, it is important to note that our design, which focused on individual differences, cannot address causation in developmental change. There are some experimental studies providing evidence that directing attention influences spatial skills, consistent with a causal role of attention. For instance, studies show that directing adults' attention through implicitly guiding eye-gaze can improve spatial recall (Bailey, McNamara, Costello, Sridharan, & Grimm, 2012). Additionally, making aspects of a task space more salient by enlarging a spatial feature in the space can improve children's use of the feature to reorient (Learmonth, Newcombe, & Huttenlocher, 2001). Evidence suggests this relation is not unidirectional, as experience with spatial activities may improve children's selective attention. Specifically, children who play more spatial games tend to perform better on spatial tasks (e.g., Jirout & Newcombe, 2015), suggesting that they may learn to attend to relevant information through spatial play.

Evidence for co-developing skills can be found across multiple domains, but the implication of correlated abilities can be difficult to interpret (van der Maas et al., 2006). One common interpretation is that one cognitive process is used to support another. However, it is also possible that cognitive systems interact reciprocally in typical experience, and these interactions lead to stronger associations even when one ability does not depend on another (see van der Maas et al., 2006, for discussion). We suggest that language is not necessarily correlated with spatial cognition because children use language to support this performance, but this does not preclude language from influencing spatial development more generally. Language development could reciprocally be interacting with both spatial cognition and selective attention, even if these skills are separable and not dependent on each other. Such interactions can be difficult to measure empirically, and cannot be detected in the design employed here. Rather it requires careful longitudinal designs that measure the relevant constructs reliably across developmental time points, as well as comprehensive theories explaining how processes interact over time.

An elegant example of such design and theory can be found in the domain of word learning. Smith and colleagues demonstrated how the process of learning words interacts with selective attention to object features, which facilitates further acceleration in children's word learning (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Specifically, Smith et al. showed that as toddlers learn labels for objects they begin to associate object names with their shapes, as most words in their vocabularies follow this regularity. This association causes children to attend to shapes when hearing new object names, which allows them to learn nouns faster. According to this theory, specific instances of noun-shape associations help toddlers learn a general characteristic (shape) to attend to while learning words. The structure of a toddler's vocabulary predicts their attention to shape while learning new words (Perry, Axelsson, & Horst, 2016), not because they are using specific words they know, but because the process of learning words has influenced their selective attention abilities.

In the domain of spatial development, similar interactions among cognitive processes could underlie the relations focused on in the current investigation. We currently do not know enough about the underlying skills' developmental trajectories to posit specific inter-relations, but we can offer some possibilities for further investigation. One is that children with good selective attention learn spatial words better because they can more easily determine the relevant dimensions involved (e.g., how to differentiate "by" from "between"). Knowledge of those words could then help children learn more about spatial layouts when hearing descriptions as well as lead them to visual explore spatial scenes in non-verbal contexts more effectively.

Alternatively, similar to Smith et al. (2002), word learning could be primary, with children's acquisition of spatial words helping them learn to attend to the more subtle visual features that differentiate more complex words (e.g., intermediate distances, direction, implicit orientations). Better attention to spatial relations could lead children to engage in spatial activities more, which could help improve their spatial skills more generally. The dynamic nature of experience over development suggests that the associations between spatial skills, selective attention, and language likely arise through mutual interaction (van der Maas et al., 2006). In the future, it will be important to conduct careful longitudinal investigations and training studies, similar to those by Smith and colleagues, to gain a clearer picture of how these factors interact to support developmental change.

These examples, building from the current results, suggest that selective attention is integrally related to language and spatial development. This adds to a growing body of research showing that selective attention is foundational to higher-level cognition (Shipstead, Harrison, & Engle, 2016) and academic success (see Stevens & Bavelier, 2012, for review). For example, one study on reading showed that connections between the visual word form area (implicated in reading) and frontal-parietal regions associated with selective attention grow stronger with age and reading experience (Vogel, Miezin, Petersen, & Schlaggar, 2012). Additionally, training adults' selective attention through video games can improve spatial and geometry skills (Novak & Tassell, 2015) and reduce sex differences in useful field of view and mental rotation tasks (Feng, Spence, & Pratt, 2007). These results suggest that selective attention could play an important role across a range of tasks, and may lay the foundation for connecting spatial skills to later STEM achievement.

In conclusion, the current study highlights the importance of considering basic attentional mechanisms in supporting young children's spatial cognition and accounting for effects previously attributed to verbal encoding. Further research will be necessary to tease apart potential causal influences within selective attention, language, and spatial cognition over development. Uncovering such causal processes will facilitate the design of effective interventions aimed at enhancing spatial skills and STEM achievement.

Acknowledgments

Thanks to the families who participated in this research, as well as the research assistants who aided in data collection and coding. Funding for participant recruitment was provided by Waisman Intellectual and Developmental Disabilities Research Center grant (P30HD03352).

References

Bailey, R., McNamara, A., Costello, A., Sridharan, S., Grimm, C. Proceedings of the Symposium on Eye Tracking Research and Applications. ACM; 2012. Impact of subtle gaze direction on short-term spatial information recall; p. 67-74.

Blumberg FC, Torenberg M, Randall JD. The relationship between preschoolers' selective attention and memory for location strategies. Cognitive Development. 2005; 20(2):242–255. https://doi.org/10.1016/j.cogdev.2005.04.006.

Clark EV, Hecht BF. Comprehension, production, and language acquisition. Annual Review of Psychology. 1983; 34(1):325–349. https://doi.org/10.1146/annurev.ps.34.020183.001545.

DeMarie-Dreblow D, Miller PH. The development of children's strategies for selective attention: Evidence for a transitional period. Child Development. 1988; 59(6):1504–1513. https://doi.org/10.1111/j.1467-8624.1988.tb03678.x. [PubMed: 3208562]

Dessalegn B, Landau B. More than meets the eye: the role of language in binding and maintaining feature conjunctions. Psychological Science. 2008; 19:189–195. https://doi.org/10.1111/j.1467-9280.2008.02066.x. [PubMed: 18271868]

Dessalegn B, Landau B. Interaction between language and vision: it's momentary, abstract, and it develops. Cognition. 2013; 127:331–344. https://doi.org/10.1016/j.cognition.2013.02.003. [PubMed: 23545385]

Dunn, LM., Dunn, DM. Peabody Picture Vocabulary Test, Fourth edition: Manual. Minneapolis, MN: Pearson Assessments; 2007.

Farran EK, O'Leary B. Children's ability to bind and maintain colour–location conjunctions: the effect of spatial language cues. Journal of Cognitive Psychology. 2016; 28(1):44–51. https://doi.org/10.1080/20445911.2015.1092980.

Feng J, Spence I, Pratt J. Playing an action video game reduces gender differences in spatial cognition. Psychological Science. 2007; 18(10):850–855. https://doi.org/10.1111/j.1467-9280.2007.01990.x. [PubMed: 17894600]

Fox, J. Regression diagnostics. Newbury Park, Calif: Sage Publications; 1991.

Hermer-Vazquez L, Moffet A, Munkholm P. Language, space, and the development of cognitive flexibility in humans: the case of two spatial memory tasks. Cognition. 2001; 79:263–299. https://doi.org/10.1016/S0010-0277(00)00120-7. [PubMed: 11165214]

Huttenlocher, J., Levine, SC. Primary test of cognitive skills. CGTB/McGraw-Hill; 1990.

Jirout JJ, Newcombe NS. Building blocks for developing spatial skills: Evidence from a large, representative U.S. sample. Psychological Science. 2015; 26(3):302–310. https://doi.org/10.1177/0956797614563338. [PubMed: 25626442]

Lauer JE, Lourenco SF. Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. Psychological Science. 2016; 27(10):1291–1298. https://doi.org/10.1177/0956797616655977. [PubMed: 27528464]

Learmonth AE, Newcombe NS, Huttenlocher J. Toddlers' use of metric information and landmarks to reorient. Journal of Experimental Child Psychology. 2001; 80:225–244. https://doi.org/10.1006/jecp.2001.2635. [PubMed: 11583524]

Levine SC, Huttenlocher J, Taylor A, Langrock A. Early sex differences in spatial skill. Developmental Psychology. 1999; 35:940. https://doi.org/10.1037/0012-1649.35.4.940. [PubMed: 10442863]

Linn MC, Petersen AC. Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development. 1985; 56(6):1479. https://doi.org/10.2307/1130467. [PubMed: 4075870]

Loewenstein J, Gentner D. Relational language and the development of relational mapping. Cognitive Psychology. 2005; 50:315–353. https://doi.org/10.1016/j.cogpsych.2004.09.004. [PubMed: 15893523]

MacKinnon DP, Krull JL, Lockwood CM. Equivalence of the Mediation, Confounding and Suppression Effect. Prevention Science. 2000; 1(4):173–181. https://doi.org/10.1023/A:1026595011371. [PubMed: 11523746]

Miller HE, Patterson AR, Simmering VR. Language supports young children's use of spatial relations to remember locations. Cognition. 2016; 150:170–180. https://doi.org/doi:10.1016/j.cognition.2016.02.006. [PubMed: 26896902]

Miller, HE., Simmering, VR. Proceedings of the 38th Annual Conference of the Cognitive Science Society. Philadelphia, PA: Cognitive Science Society; 2016. General mechanisms underlying language and spatial cognitive development.

Miller, HE., Vlach, HA., Simmering, VR. Producing spatial words is not enough: Understanding the relation between language and spatial cognition. Child Development. 2017. https://doi.org/10.1111/cdev.12664

Newcombe, NS., Ratliff, KR. Explaining the development of spatial reorientation: Modularity-plus-language versus the emergence of adaptive combination. In: Plumert, JM., Spencer, JP., editors. The emerging spatial mind. Oxford University Press; 2007. p. 53-76.

Novak E, Tassell J. Using video game play to improve education-majors' mathematical performance: An experimental study. Computers in Human Behavior. 2015; 53:124–130. https://doi.org/10.1016/j.chb.2015.07.001.

Perry LK, Axelsson EL, Horst JS. Learning What to Remember: Vocabulary Knowledge and Children's Memory for Object Names and Features: Brief Report. Infant and Child Development. 2016; 25(4):247–258. https://doi.org/10.1002/icd.1933.

Pickering SJ. The development of visuo-spatial working memory. Memory. 2001; 9:423–432.

Plumert JM. Young children's ability to detect ambiguity in descriptions of location. Cognitive Development. 1996; 11(3):375–396. https://doi.org/10.1016/S0885-2014(96)90010-6.

Plumert JM, Nichols-Whitehead P. Developmental differences in preferences for using color, size, and location information to disambiguate hiding places. Journal of Cognition and Development. 2007; 8:427–454.

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

Pruden SM, Levine SC. Parents' spatial language mediates a sex difference in preschoolers' spatial-language use. Psychological Science. 2017; 28(11):1583–1596. https://doi.org/10.1177/0956797617711968. [PubMed: 28880726]

Pruden SM, Levine SC, Huttenlocher J. Children's spatial thinking: Does talk about the spatial world matter? Developmental Science. 2011; 14:1417–1430. https://doi.org/10.1111/j.1467-7687.2011.01088.x. [PubMed: 22010900]

Riggs KJ, Simpson A, Potts T. The development of visual short-term memory for multifeature items during middle childhood. Journal of Experimental Child Psychology. 2011; 108:802–809. http://dx.doi.org/10.1016/j.jecp.2010.11.006. [PubMed: 21257180]

Roderer T, Krebs S, Schmid C, Roebers CM. The Role of Executive Control of Attention and Selective Encoding for Preschoolers' Learning: Selective Encoding and Executive Control of Attention. Infant and Child Development. 2012; 21(2):146–159. https://doi.org/10.1002/icd.738.

Shipstead Z, Harrison TL, Engle RW. Working memory capacity and fluid intelligence: maintenance and disengagement. Perspectives on Psychological Science. 2016; 11(6):771–799. https://doi.org/10.1177/1745691616650647. [PubMed: 27899724]

Shusterman A, Lee S, Spelke ES. Cognitive effects of language on human navigation. Cognition. 2011; 120:186–201. https://doi.org/10.1016/j.cognition.2011.04.004. [PubMed: 21665199]

Shusterman, A., Spelke, E. Language and the development of spatial reasoning. In: Carruthers, P.Laurence, S., Stich, S., editors. The innate mind: Structure and concepts. Oxford, UK: Oxford University Press; 2005. p. 89-106.

Simmering, VR., Wood, CM. The importance of real-time stability for visual working memory performance: Young children's feature binding can be improved through perceptual structure. Developmental Psychology. 2017. https://doi.org/10.1037/dev0000358

Simms, N., Gentner, D. 29th Annual Conference of the Cognitive Science Society. Washingtion, D.C.: Lawrence Erlbaum Associates; 2008. Spatial language and landmark use: Can 3-, 4-, and 5-year-olds find the middle?; p. 191-196.

Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, Samuelson L. Object name learning provides on-the-job training for attention. Psychological Science. 2002; 13:13–19. http://dx.doi.org/10.1111/1467-9280.00403. [PubMed: 11892773]

Stevens C, Bavelier D. The role of selective attention on academic foundations: A cognitive neuroscience perspective. Developmental Cognitive Neuroscience. 2012; 2:S30–S48. https://doi.org/10.1016/j.dcn.2011.11.001. [PubMed: 22682909]

van der Maas HLJ, Dolan CV, Grasman RPP, Wicherts JM, Huizenga HM, Raijmakers MEJ. A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. Psychological Review. 2006; 113(4):842–861. [PubMed: 17014305]

Vasilyeva, M., Lourenco, SF. Spatial development. In: Lerner, RM.Lamb, ME., Freund, AM., editors. The Handbook of Life-Span Development. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2010. Retrieved from http://doi.wiley.com/10.1002/9780470880166.hlsd001020

Verdine BN, Golinkoff RM, Hirsh-Pasek K, Newcombe NS, Filipowicz AT, Chang A. Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. Child Development. 2014; 85:1062–1076. https://doi.org/10.1111/cdev.12165. [PubMed: 24112041]

Vogel AC, Miezin FM, Petersen SE, Schlaggar BL. The putative visual word form area is functionally connected to the dorsal attention network. Cerebral Cortex. 2012; 22(3):537–549. https://doi.org/10.1093/cercor/bhr100. [PubMed: 21690259]

Wai J, Lubinski D, Benbow CP. Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. Journal of Educational Psychology. 2009; 101:817–835. http://dx.doi.org/10.1037/a0016127.

Highlights

- Prior research suggests that spatial language supports spatial skill development
- The hypothesis that selective attention account for this relation is tested further
- Attention accounts for more variance in spatial skills than adaptive language use
- Adaptive language use reflects both attention and task-related language production
- Selective attention could be a domain-general mechanism in spatial development

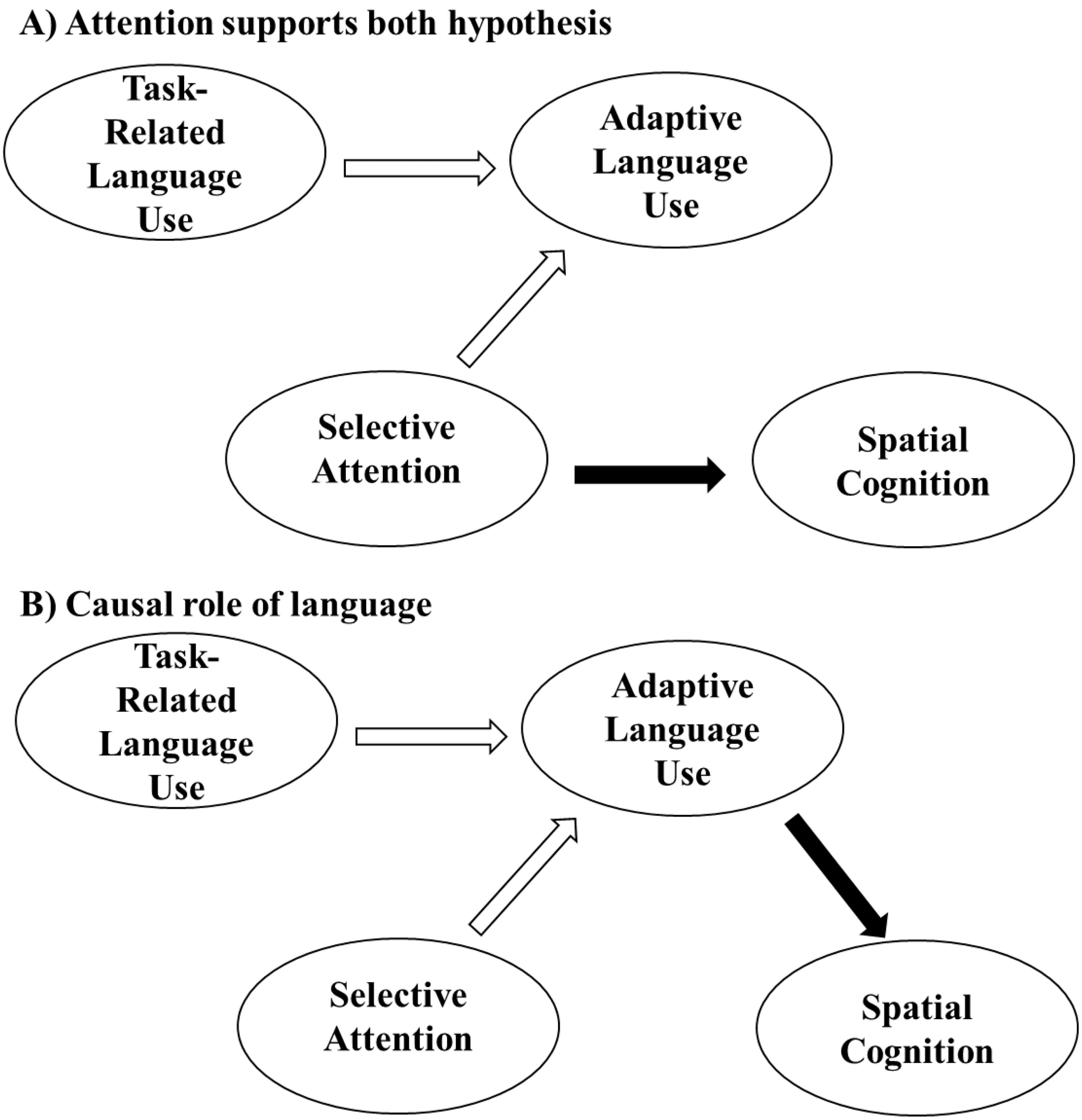

Figure 1.

A) Represents hypothesized causal pathways for hypothesis that attention supports both spatial skills and task-relevant language use. B) Represents hypothesized causal pathway for alternative hypothesis positing causal role of language.

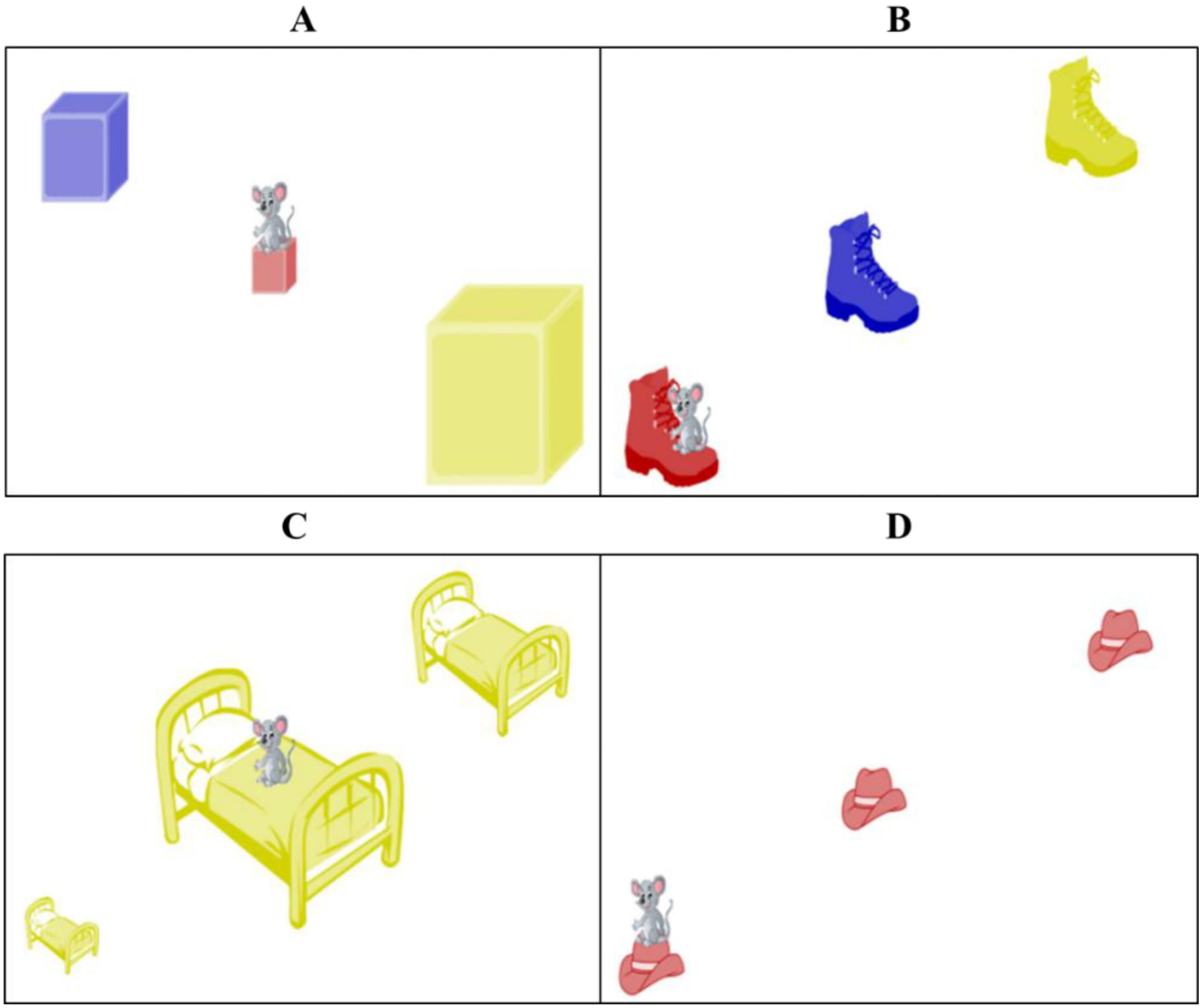

Figure 2.
Examples of the four scene types in the spatial scene description task: A) a three cue scene with color, size, and location cues; B) a two cue scene with color and location cues; C) a two cue scene with size and location cues; and D) a one cue scene with only location cues.

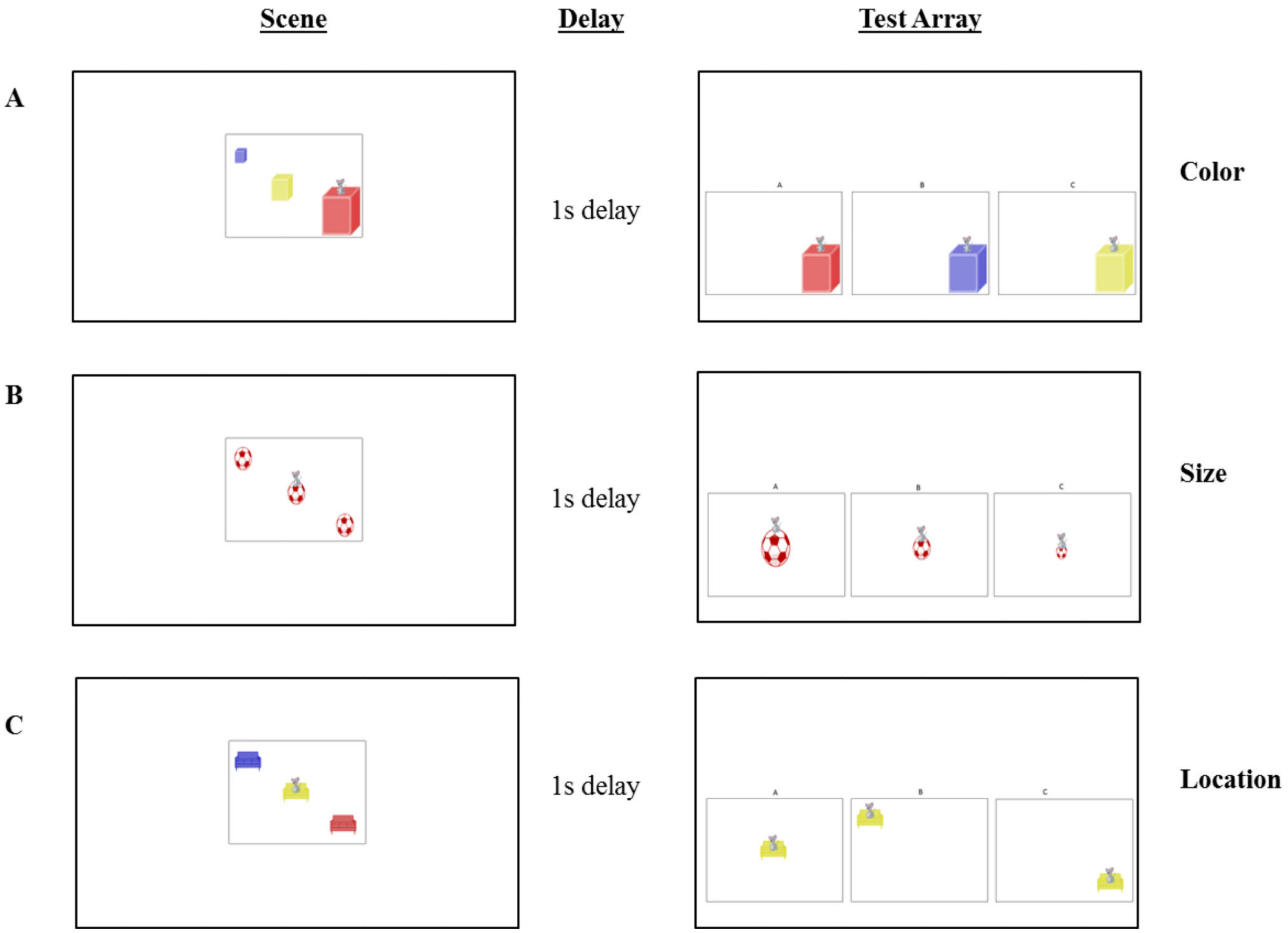

Figure 3.
Sample trials of the spatial scene attention task: A) a 3-cue scene with color probe, B) a 1-cue scene with size probed, and C) a 2-cue-color scene with location probed.

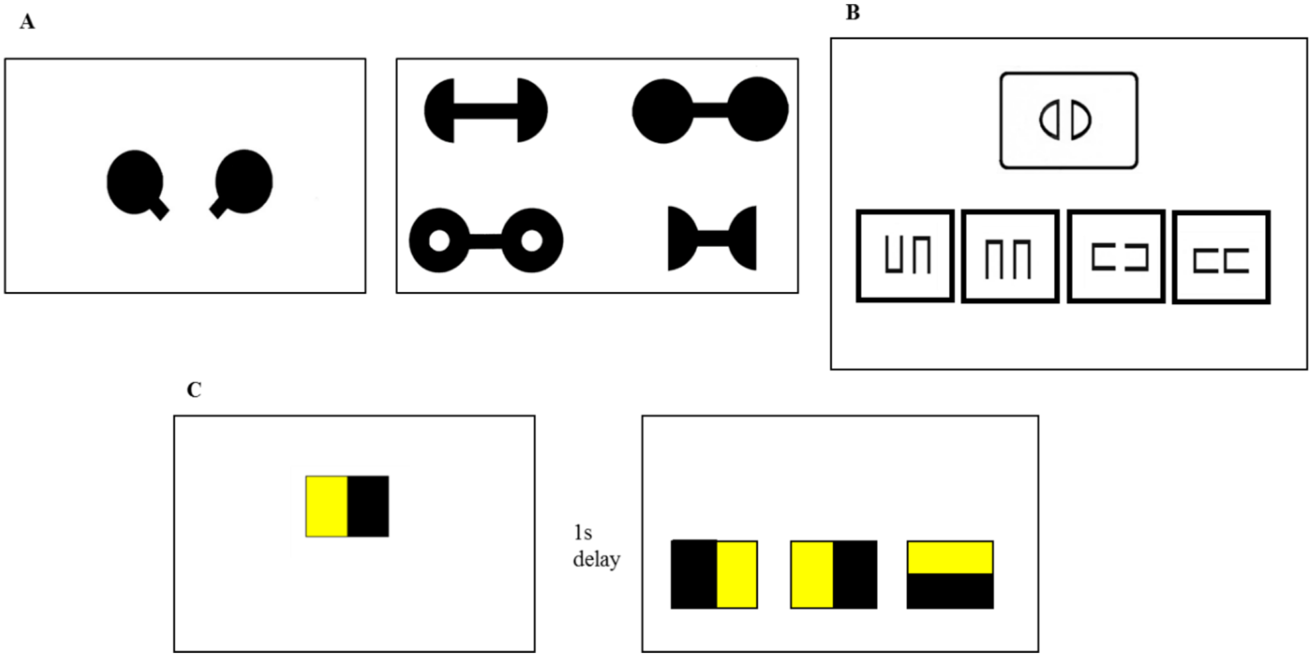

Figure 4.

Sample trials of spatial tasks. A) In the Children’s Mental Transformation Task, children were asked to “select the shape that the pieces make.” B) In the Spatial Analogies Task, children were asked to “select the picture that goes best with that picture (target picture).” C) In the Feature Binding Task, children saw the target square and were asked after a 1s delay, to “select the square that is the same as the one you just saw.”

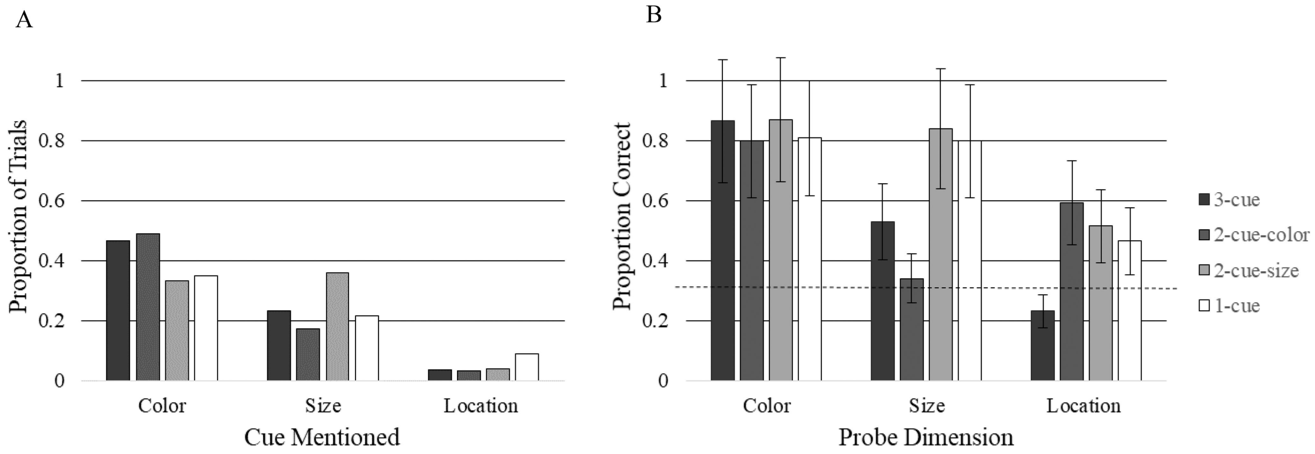

Figure 5.

A) Proportion of cues mentioned in spatial scene description task from Experiment 1 by cue type and trial type. B) Proportion correct on spatial scene attention task from Experiment 1 for each probe dimension by trial type. Error bars show 95% confidence intervals of the mean. In the attention task, chance = .33 (dashed horizontal line).

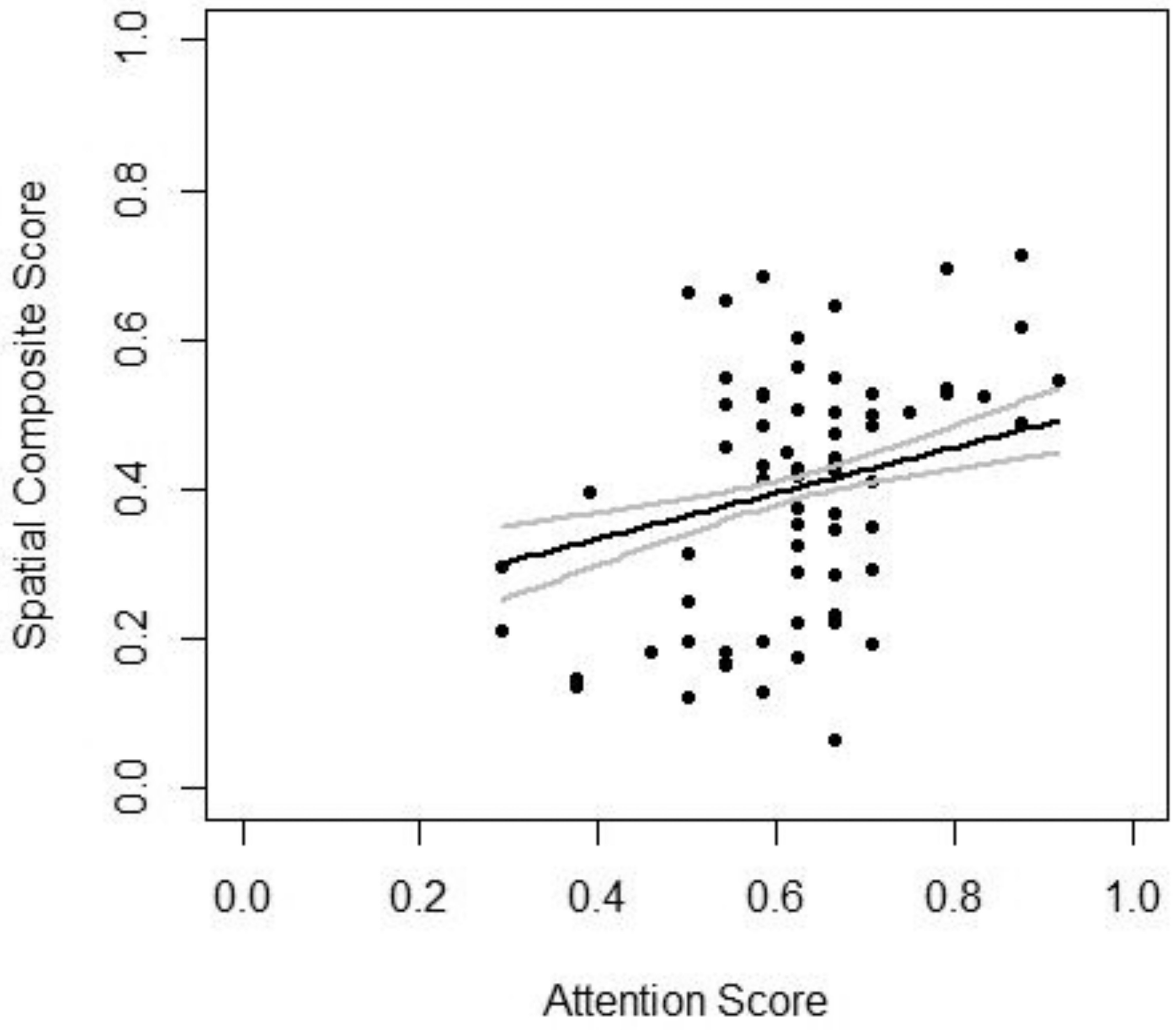

Figure 6.

Relations between children's spatial composite score and attention score, while controlling for age, gender, PPVT-IV score, and adaptation score (black lines). The error bars represent ± 1 standard error for point estimates from the regression model (grey lines). Chance performance for spatial composite score was 0, and was .33 for the attention task.

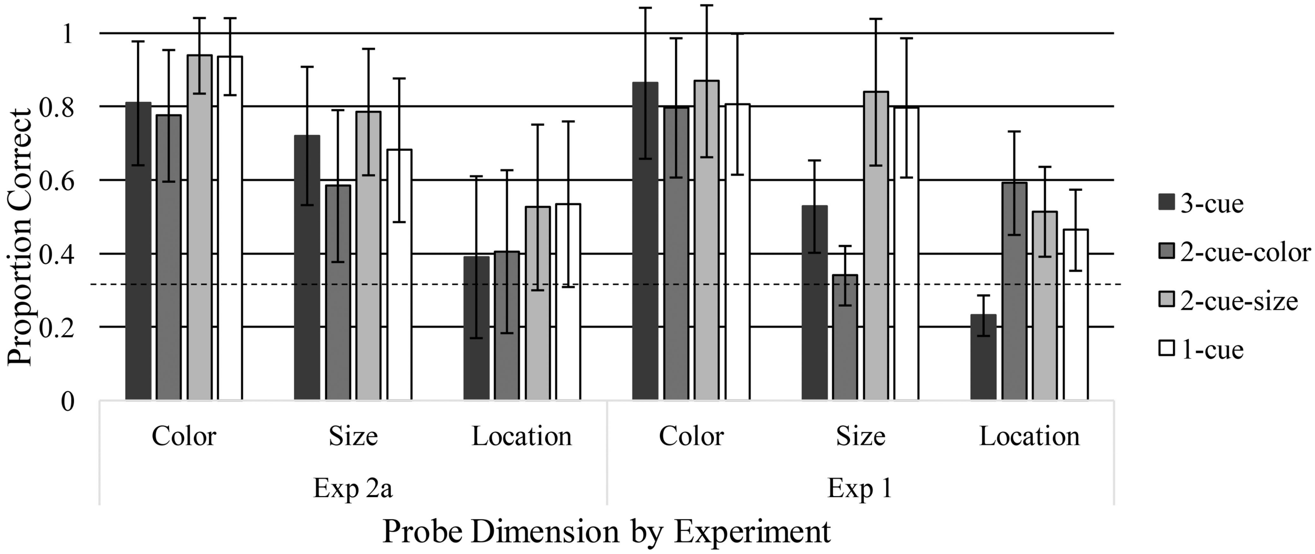

Figure 7.

Proportion correct on spatial scene attention task for each probe dimension by trial type and Experiment (2a, 1). Error bars show 95% confidence intervals of the mean. Chance = .33 (dashed horizontal line).

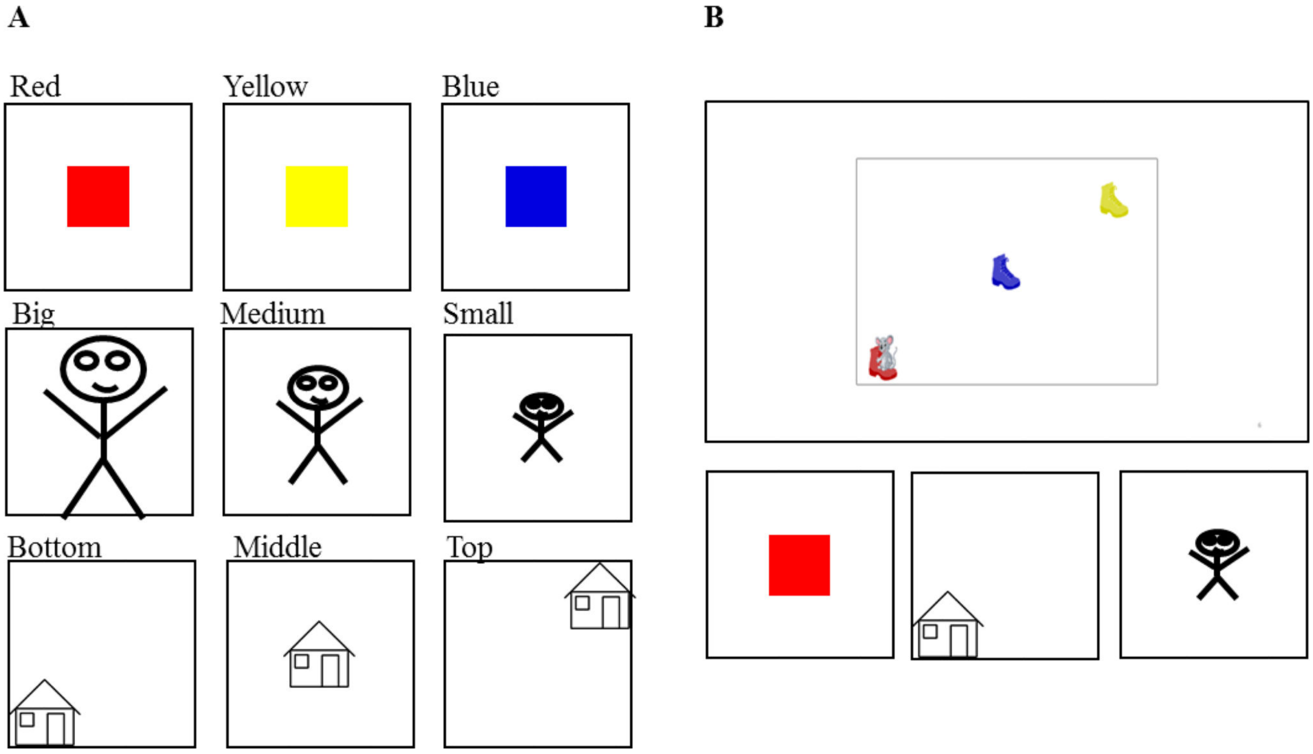

Figure 8.

Spatial scene cue-selection task. A) Color, size, and location cue cards, with the corresponding words the experimenter provided, used in task. B) Sample trial from task with image of spatial scene and corresponding cue cards.

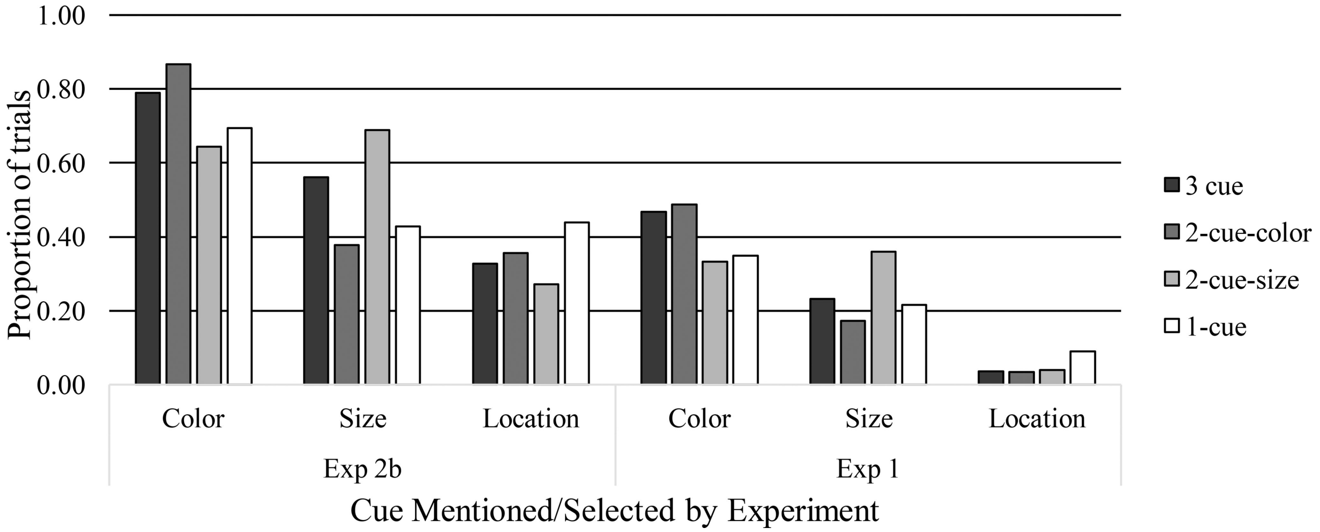

Figure 9.

Proportion of cues mentioned in spatial scene cue-selection (Experiment 2b) and description task (Experiment 1) by cue type and trial type.

Table 1

Descriptive Statistics of Measures from Experiment 1

Measures	**Mean**	**Standard Deviation**	**95% Confidence Interval**	**Observed Range**
Spatial Composite Score	0.40	0.17	0.36, 0.44	0.07 to 0.71
Age	4.51	0.40	4.42, 4.61	3.83 to 5.19
PPVT-IV Standard Score	120.65	15.61	116.94, 124.36	75.00 to 148.00
Adaptation Score	0.29	0.39	0.19, 0.38	−0.17 to 1.56
Attention Score	0.62	0.13	0.59, 0.65	0.28 to 0.92
Task-Related Production	0.71	0.58	0.57, 0.84	0.00 to 1.92
Mental Transformation	0.59	0.21	0.57,0.61	0.10 to 0.90
Spatial Analogies	0.39	0.18	0.37,0.40	0.08 to 0.77
Feature Binding	0.70	0.17	0.69,0.72	0.13 to 1.00

Note. PPVT-IV= Peabody Picture Vocabulary Test, 4th edition. Spatial Composite scores were proportion correct, normalized for differing levels of chance across tasks (see text for details), resulting in a chance level of 0. The PPVT-IV scores were standardized, with possible scores ranging from 20 to 160. Spatial Vocabulary was calculated as the proportion of terms caregivers checked (out of 80). The Adaptation score could range from −2 to 3 (see text for description). The Attention score was overall proportion correct, chance is .33. Task-Related Production Score was calculated as the average number of unique color, size, and location terms children used per trial. Mental Transformation, Spatial Analogies, and Feature Binding scores were calculating by taking the average raw scores. Chance was 0.25 for Mental Transformation and Spatial Analogies and 0.33 for Feature Binding.

Table 2

Pearson’s Correlations among Measures in Experiment 1

Measures	1	2	3	4	5	6	7	8	9
1. Spatial Composite Score									
2. Age	$.45^{***}$								
3. Gender	−.10	−.02							
4. PPVT-IV Standard Score	$.41^{***}$	−.03	.02						
5. Adaptation Score	$.34^{**}$	$.34^{**}$	$.39^{***}$	.08					
6. Attention Score	$.45^{***}$	$.39^{***}$	.08	.10	$.43^{***}$				
7. Task-Related Production	$.30^{*}$	$.28^{*}$	.06	.06	$.56^{***}$	$.24^{*}$			
8. Mental Transformation	$.71^{***}$	$.39^{***}$	−.03	$.43^{***}$	$.28^{*}$	$.29^{*}$	$.29^{*}$		
9. Spatial Analogies	$.75^{***}$	$.29^{*}$	−.09	$.37^{**}$	$.22^{\dagger}$	$.36^{**}$	$.22^{\dagger}$	$.44^{***}$	
10. Feature Binding	$.46^{***}$	.18	−.08	−.02	.15	$.23^{\dagger}$	.06	−.13	.02

Note.

$^{\dagger}p < .071$,

$^{*}p < .05$,

$^{**}p < .01$,

$^{***}p < .01$. Gender was contrast coded (−.5 = female, .5 = male). PPVT-IV= Peabody Picture Vocabulary Test 4th edition.

Table 3

Regression Analysis Predicting Adaptation Score

	b	SE_b	β	R^2	*F*	*p*
Predictors						
				.53	13.74	<.001
Age	0.13	0.10	0.13			
Gender	0.26	0.07	0.34***			
PPVT-IV Standard Score	0.00	0.00	0.03			
Task-Related Production	0.29	0.06	0.44***			
Attention Score	0.74	0.30	0.24 *			

Note.

* $p < .05$, $p <.001$. Gender was contrast coded (−.5 = female, .5 = male). PPVT-IV= Peabody Picture Vocabulary Test 4th edition.

Table 4

Hierarchical Regression Analysis Predicting Spatial Composite Score

	b	SE_b	β	R^2	ΔR^2	ΔF	p
Predictors							
Step 1: Adaptation Score				.44	.44	12.51	<.001
Age	0.16	0.04	0.37***				
Gender	−0.07	0.03	-0.20^{t}				
PPVT-IV Standard Score	0.00	0.00	0.40***				
Adaptation Score	0.11	0.05	0.26^{*}				
Step 2: Attention Score				.48	.04	4.74	.033
Age	0.13	0.04	0.31^{**}				
Gender	−0.06	0.03	-0.19^{t}				
PPVT-IV Standard Score	0.00	0.00	0.38***				
Adaptation Score	0.08	0.05	0.18				
Attention Score	0.31	0.14	0.23^{*}				

Note.

[t] $p < .058$,

[*] $p < .05$,

[**] $p < .01$, $p < .001$. Gender was contrast coded (−.5 = female, .5 = male). PPVT-IV= Peabody Picture Vocabulary Test 4th edition.