Available online at www.sciencedirect.com
SCIENCE DIRECT®
Cognitive Psychology 50 (2005) 315–353
Cognitive Psychology
www.elsevier.com/locate/cogpsych

Relational language and the development of relational mapping ☆

Jeffrey Loewenstein [a], Dedre Gentner [b,*]

[a] *University of Texas at Austin, USA*
[b] *Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208-2710, USA*

Accepted 22 September 2004
Available online 11 January 2005

Abstract

We test the claim that learning and using language for spatial relations can influence spatial representation and reasoning. Preschool children were given a mapping task in which they were asked to find a "winner" placed in a three-tiered box after seeing one placed in a virtually identical box. The correct choice was determined by finding the corresponding relative location in the test box, making it a difficult task for preschool children. We found that hearing language for spatial relations facilitated children's mapping performance. We found effects at younger ages on easier tasks (Experiments 1 and 2) and at older ages on harder tasks (Experiment 3). The effects of spatial relational language differed predictably according to the semantics of the terms children heard (Experiment 4). Finally, the effects of spatial language were maintained over time (Experiment 5): children given one initial exposure to the spatial terms maintained their advantage over baseline children when they again carried out the mapping task 2 days later, with no further exposure to the spatial terms. The evidence is consistent with the explanation that language bolsters children's spatial encodings, which in turn supports their mapping performance.

☆ This research was supported by NSF Grant SBR-9720313/5-27481, NSF-LIS Grant SBR-9511757, and NSF-ROLE Grant 21001/REC-0087516 to Dedre Gentner. We thank Shelli Feist, David Uttal, Phil Wolff, and the Analogy and Similarity group at Northwestern University for helpful discussions of these issues, as well as Rochel Gelman and the anonymous Cognitive Psychology reviewers for their insights. We thank Michelle Osmondson and Barbara Hung for their assistance in carrying out the studies, and the many parents and children who participated.
* Corresponding author. Fax: +1 847 491 7859.
E-mail address: gentner@northwestern.edu (D. Gentner).

0010-0285/$ - see front matter
doi:10.1016/j.cogpsych.2004.09.004

1. Introduction

An appreciation of relational similarity is a hallmark of human cognition. We readily notice commonalities such as the match in configuration between a map and a city, or the match in predatory behavior between a shark and a tiger. Relational similarity is central to analogy (Gentner, 1983, 2003), and is important in inductive inferencing (Holland, Holyoak, Nisbett, & Thagard, 1986), and categorization (Ramscar & Pain, 1996). Many everyday terms denote categories based on relational similarities (Gentner & Kurtz, in press). For example, to learn and apply a category like 'predator' or 'obstacle' requires perceiving a common relational pattern—such as "lives by consuming other animals," or "blocks the attainment of some entity's goal." How do children come to appreciate these non-obvious commonalities? In this research we explore the possibility that language for relations promotes children's understanding of relational similarities. Our claim is that relational language fosters the development of representational structures that facilitate mental processing—that is, that relational language provides tools for thought.

In proposing that relational language influences children's thinking, we are not proposing a strong Whorfian hypothesis of linguistic determinism. The issue of whether and how language may influences cognition is currently being revisited and re-debated (Hunt & Agnoli, 1991; Levinson, Kita, Haun, & Rasch, 2002; Li & Gleitman, 2002; Slobin, 2003), and subtler possibilities are now being discussed. Gentner and Goldin-Meadow (2003a, 2003b) differentiated the language and thought question into three categories. The *language as lens* view is the classic Whorfian hypothesis that the grammatical structure of a language shapes its speakers' perception of the world. At the other extreme, the *language as category shifter* view maintains that conceptual categories are universal, but language can influence their boundaries. In the *language as tool kit* view, language provides concepts and strategies that augment, but do not supplant, other methods of representation and reasoning. This view is related to Vygotsky's (1962) claim that language is instrumental in learning to direct mental processes, but differs in emphasizing specific semantic and grammatical devices. Our approach fits under this third view. We suggest that relational language provides tools for extracting and formulating relational representations (Gentner & Loewenstein, 2002), and thereby potentiates analogy and other processes that operate over relational structure (Gentner, 2003). Of course, language is not the only medium that fosters relational learning—maps, diagrams, and mathematics, among others, can support relational cognition. But language is a superb instrument for naming and expressing relational structures, and acquiring language is a widespread conventional means for such learning.

1.1. The relational shift

The apprehension of relational commonalities is not immediate in learning and development. Children at first rely on overall similarity or on object-level commonalities, then shift to appreciating relational similarities (Gentner & Rattermann, 1991; see also Halford, 1993). For example, Chen, Sanchez, and Campbell (1997) found that 10-month-old infants who learned to pull on a cloth to reach a toy could transfer this pulling relation to a new situation only if the new situation was highly similar to the initial situation. By 1;1 (years; months), infants were able to transfer with less concrete similarity. As another example, Smith (1984) investigated the development of similarity with a follow-the-leader task in which two experimenters (E1 and E2) in turn chose objects from sets of toys, and the child was told to choose in the same way. The results showed a clear order of emergence of similarity kinds among 2- to 4-year-olds. Direct object matches were grasped very early—for example, if E1 and E2 each took a green plane, 2-year-olds readily took a green plane. Two-year-old children could also notice and apply common object properties (e.g., E1: blue boat, E2: blue house; child: blue car). However, their performance dropped sharply on trials involving common dimensional relations (e.g., E1: green house, green car; E2: blue boat, blue bridge; child: two (different) yellow objects). Another indication of the difficulty of relational concepts is that children are relatively slow to grasp the meanings of relational nouns, and often initially interpret them as object names. For example, Hall and Waxman (1993) found that $3\frac{1}{2}$-year-olds had difficulty in learning novel relational nouns denoting concepts like *passenger.* Even when they were explicitly told (for example) "This one is a *blicket* BECAUSE IT IS RIDING IN A CAR," children tended to interpret the novel noun as referring to the object category.

The relational shift can also be seen in the spatial domain. Researchers have noted that children understand element-to-element correspondences before they understand spatial relational correspondences (e.g., Bluestein & Acredolo, 1979; Presson, 1982). Liben (1998) has described a comparable pattern in map understanding in which object-based ("representational") correspondences are grasped before relation-based ("geometric") correspondences. This shift from object similarity to relational similarity also occurs in spatial mapping tasks, in which children watch a toy being hidden in one model room and then are asked to find a similar toy hidden in the same place in a second model room (DeLoache, 1995). For example, Blades and Cooke (1994) found evidence of this developmental sequence using model rooms that each contained two identical items, along with two other (uniquely matching) objects. When toys were hidden at a unique object, the 3-year-old children searched successfully at the corresponding object. But for toys hidden at one of the identical objects, they guessed randomly between the two identical objects in the second model. Older children were less strongly reliant on object matches: four-year-olds could search correctly within the identical pairs by using common spatial relations to disambiguate the object matches (see also Loewenstein & Gentner, 2001).

1.2. The role of language

As suggested by the above discussion, the relational shift occurs at different times for different tasks and domains. Further, it is influenced by domain expertise. For example, Gobbo and Chi (1986) found that 7-year-old dinosaur experts were better able than age-matched novices to draw appropriate inferences about new dinosaurs and more often compared them with relationally similar dinosaurs. Thus the shift to relational matching appears to be driven in part by gains in domain knowledge (Brown, 1989; Gentner & Rattermann, 1991), as well as possibly by maturational changes in processing capacity (Halford, 1993). In this research, we explore one route to learning domain relations: the learning and application of relational language.

There is considerable evidence that common language can serve as an invitation to seek common concepts. Word-learning studies have repeatedly shown that when children are taught a new object term, they tend to assume that the word applies to things of like kind (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Waxman & Gelman, 1986). For example, Markman and Hutchinson (1984) gave 2- to 3-year-olds a standard such as a spider along with two alternatives: a web and a fly. When asked what the spider goes with, children often chose the thematically related alternative, the web. But when the spider was named with a new word, and children were asked to "find the other *dax*" they mostly chose the like item—the fly.

These kinds of findings suggest that language can invite children to form and use certain kinds of representations. To date, research on the possible conceptual effects of language has concentrated on object terms. However, it seems at least equally plausible that language for relations might foster attending to, and encoding, particular relations (Logan & Sadler, 1996; Regier & Carlson-Radvansky, 2001). If so, then once relational terms have been acquired, hearing relational language might facilitate encoding relations in ways consistent with the semantics of the terms.

Relational terms—and in particular, spatial relational terms—are an important arena for several reasons. First, as Gentner (1981, 1982; Gentner & Boroditsky, 2001; Gentner & Goldin-Meadow, 2003a, 2003b) suggests, relational terms such as verbs and prepositions may have a greater influence on how relations are represented than nouns do on how objects are represented. Because relational terms are more variable cross-linguistically than nominal terms of comparable concreteness, they allow greater scope for concomitant cognitive variation. For example, Talmy's (1983) analysis of motion verbs showed differences in the characteristic semantic patterns adopted in different languages. Second, spatial relational terms have been found to exhibit considerable variability across languages. For example, investigations of the semantics of terms related to contact, support, and inclusion (in English, '*in*' and '*on*') have revealed striking cross-linguistic variation (Bowerman & Choi, 2001; Feist, 2000). Likewise, spatial orientation terms such as left–right and north–south vary across languages, and there is cross-cultural evidence suggesting that some of these semantic variations may influence habitual spatial reasoning patterns of their speakers (for recent overviews of this literature, see Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003a, 2003b; Gumperz & Levinson, 1996; also see Li & Gleitman, 2002, for a countervailing view). A second advantage of spatial relations as a domain of inquiry is that, because they apply to the perceptual arena, it is relatively easy to devise non-linguistic tasks that can be assessed with or without the use of language. Finally, spatial relations are of interest because of their intrinsic importance and wide applicability in both spatial and non-spatial domains (Gattis, 2001; Gentner, Imai, & Boroditsky, 2002).

Some recent studies have provided suggestive evidence that the use of spatial language can influence the way in which people represent and reason about space. Hermer-Vasquez, Moffet, and Munkholm (2001) gave preschool children a spatial retrieval task in which they had to remember the location of an object hidden in one corner of a white, rectangular room with blue fabric on one wall that could serve as a reference point. They found that children's performance on the retrieval task was correlated with their ability to use the spatial language relevant to the task. Hermer-Vasquez et al. suggested that spatial language provides children with a way of communicating between two otherwise separate spatial representation systems—one specialized for geometric relations and one specialized for egocentric relations between the self and a landmark. The role of language was further underscored by the finding that adults performed poorly on the spatial task when conjointly performing a verbal shadowing task (Hermer-Vasquez, Spelke, & Katsnelson, 1999). Verbal shadowing did not appear to impair adults' abilities to make either the left–right distinction or the color distinction by themselves, but it impaired their ability to relate the two (and thus to determine the object's location). Although there has been some debate concerning the source of children's difficulties (Newcombe, 2002; Newcombe & Huttenlocher, 2000), the findings strongly suggest a role for spatial relational language in supporting spatial cognition.

The studies we will present examine effects of relational language on spatial mapping performance. Our interest is in the effects of acquiring and using spatial language within a language community. The hypothesis is that language provides representational tools with which speakers can create construals that facilitate reasoning. We used a version of DeLoache's (1995) classic spatial mapping task: children watched an item being hidden in one space and then were asked to find a matching item hidden in the corresponding place in a second space. Because of our focus on spatial language, we used as our spaces two vertically arrayed boxes (identical except for color), each with three locations at which an object could be placed (Fig. 1). The three locations within each box are readily named using either of two sets of spatial terms in English—*on*, *in*, *under* and *top*, *middle*, *bottom* —permitting us to compare performance with and without such language. The correct choice was always in the same relative position in the two boxes.

Based on prior research on the development of analogical mapping and search, we predicted that this relational mapping task would be difficult for preschool children. If hearing spatial relational language induces children to adopt specific spatial encodings, then the use of such language should facilitate mapping performance. To clarify what we mean by saying that language encourages the encoding of "specific spatial relations," consider that a child (or adult) seeing an object placed in or on the box could encode the relation rather vaguely, roughly as "It's at the box." To the extent that young children tend to use such rather general encodings, they will fail to show a differentiated spatial mapping between the boxes. Our hypothesis is that hearing spatial terms will encourage representing the spatial relations within each box—just what is needed to permit a structural alignment between the boxes. Experiment 1 tests this prediction for the two sets of spatial terms. Experiment 2 addresses alternative interpretations of our findings. In Experiment 3, we made the task

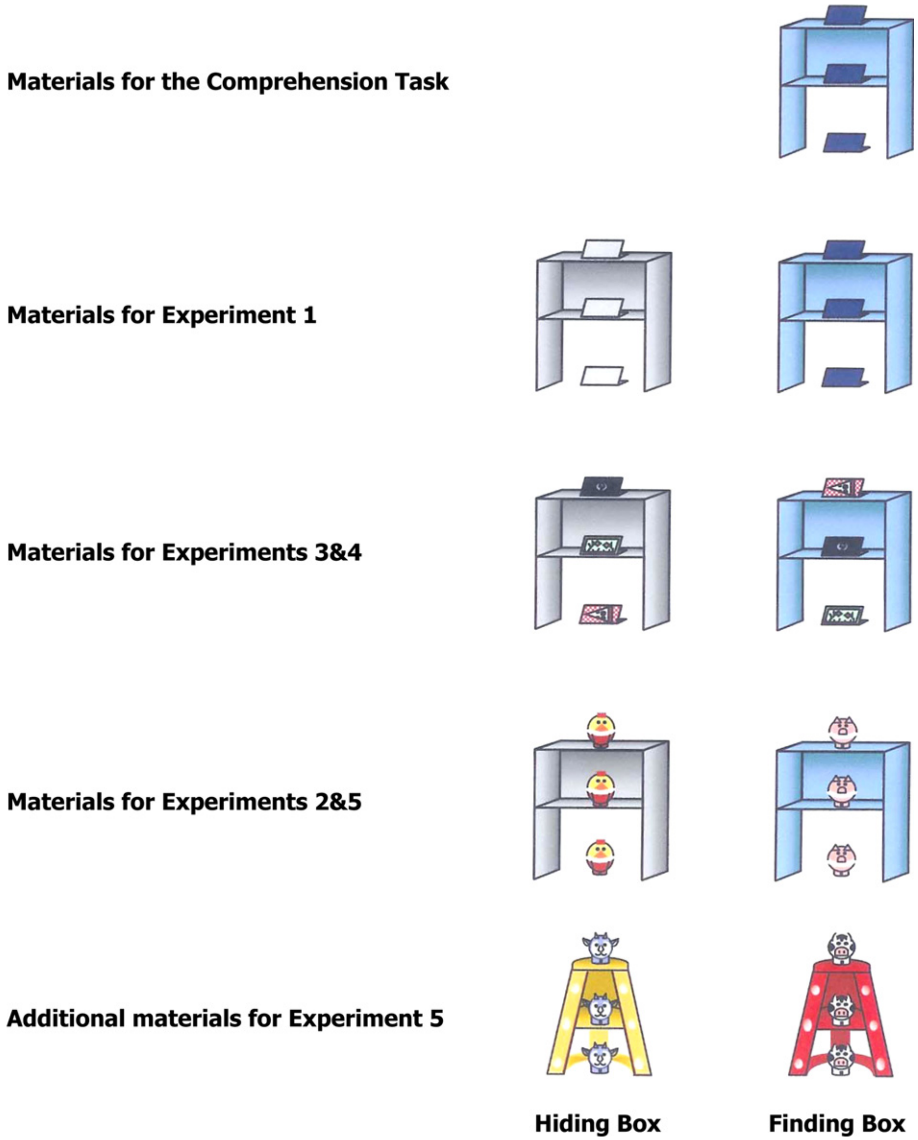

Fig. 1. Materials for the Comprehension task and all Experiments.

more complex to examine whether older children benefit from hearing spatial language. Experiment 4 tests a further prediction arising from the claim that the effects of spatial language result from inducing specific relational representations according to the semantics of the terms. To test this claim, we contrasted the two sets of spatial terms on the challenging mapping task. As demonstrated in the accompanying cognitive simulation, the representations invited by these two sets of terms should lead to different outcomes on this task. Finally, in Experiment 5 we test whether hearing spatial relational language yields a conceptual representation sufficiently durable to support children's performance 2 days later.

To summarize, our central claim is that hearing language for spatial relations will foster encoding specific spatial relations. This richer and more delineated relational encoding will support mapping on the basis of those relations. The following studies test four predictions. First, hearing language for spatial relations will facilitate children's performance in a spatial mapping task. Second, effects of spatial relational language will be seen at later ages for more difficult tasks than for easier tasks. Third, the effects of spatial relational language will be specific to the semantics of the terms used. Fourth, the effects of spatial language will persist over time.

Comprehension study. Our studies depend on children knowing the meanings of the spatial terms we use. Prior studies show that English-speaking children acquire many spatial relational terms in the preschool years, but also that children's initial understanding and use of these terms is context-specific (Clark, 1980; Kuczaj & Maratsos, 1975; Meints, Plunkett, Harris, & Dimmock, 2002). The terms relevant for our spatial mapping tasks are *on*, *in*, and *under* and *top*, *middle*, and *bottom*. Johnston (1988) summarized results from a variety of studies showing that *on*, *in*, and *under* are comprehended quite early—by about 2;6 for typical uses. Tomasello (1987) in his diary study found that *on*, *in*, and *under* were learned by early in the third year. Clark (1980) found that by 3;3, children can point to the *top* and *bottom* of vertically oriented objects. Although there is little data on children's comprehension of *middle*, there is some evidence that children comprehend *between* early in the third year (Johnston, 1988).

We tested 3-year-olds' comprehension of the spatial terms in the specific contexts used in our studies (Fig. 1). The reference object used in our studies was a box with three clearly distinguishable locations: on top, in the middle, and at the bottom (modeled after a "contextually neutral" apparatus used by Wilcox & Palermo, 1980). We showed 3;2- and 3;8-year-old children the box reference object with a card at each of the three locations, then tested their comprehension of the two word sets *on*, *in*, and *under* and *top*, *middle*, and *bottom*. We asked them to point to the card that was *on* the box (or *at the top of* the box, and so forth). Children were tested twice for each term, one word set at a time, and the order of presentation was counterbalanced. The results are summarized in Table 1.

The results accord with prior findings that 3-year-old children show high levels of comprehension for these basic spatial terms. By 3;8, children show accurate comprehension of both word sets. They were 85% "correct" for *on*, *in*, *under*, and 89%

Table 1
Comprehension of spatial terms: proportion correct identifications of location by Age and Word

Age	*n*	On	In	Under	OIU mean	Top	Middle	Bottom	TMB mean
3;2 years	20	.60	.90	.90	.80	1.0	.80	.85	.88
3;8 years	16	.84	.81	.91	.85	.94	.84	.88	.89

Note. "OIU," *On*, *In*, *Under* and "TMB," *Top*, *Middle*, *Bottom*.

correct for *top*, *middle*, *bottom*. The 3;2-year-old children were also highly accurate, with the possible exception of *on* (for which they sometimes chose the middle location). Overall, although some individual terms may be ambiguous, the results show that by 3;8 children have arrived at a stable interpretation for each word set, in our context. As the youngest age to participate in the studies that follow is 3;7, we can be reasonably confident that children will interpret the word sets as intended.

2. Experiment 1

Our central hypothesis is that spatial language might invite young children to encode spatial relations, which in turn could facilitate noticing relational similarities. In Experiment 1 we tested whether the overt use of spatial relational language would improve young children's performance in a spatial mapping task. Children were given a spatial mapping task using two boxes—a Hiding box and a Finding box—as shown in Fig. 1. Children watched as the "winner" card was placed at the Hiding box and had to find the other "winner" by searching in the corresponding place at the Finding box. To test children's ability to carry out a purely relational mapping task, all the cards had blank fronts. (The winner had a star on its back.) Thus, at the level of object similarity, each card could match any other card. The only basis for drawing correspondences was spatial relational information. For half the children, the experimenter labeled the location of the winner card with a specific spatial relational term (e.g., "I'm putting this *on* the box"). The other half of the children performed the same task with a general reference to location; they were told "I'm putting this here." We predicted that children who heard specific spatial terms would be more likely to find the winner than children not given such descriptors. We further predicted that hearing spatial language would lead to an advantage in children's ability to articulate how they knew where the winners were in response to a question from the experimenter. Previous studies have found positive correlations between children's understanding of relational language and their performance on spatial relational tasks (Hermer-Vasquez et al., 2001; Kuenne, 1946). Here we seek a more direct evidential link.

2.1. Method

2.1.1. Participants

The participants were 88 preschool children, mostly from middle and upper-middle class families in the Chicago area, recruited via home mailings. No child participated in more than one experiment. There were two age groups, averaging 3;8 years (range: 3;5–3;10 years), and 4;1 years (range: 3;11–4;4 years). Within each age group, children were randomly assigned to the baseline group ($n = 22$) or the language group ($n = 22$). The language group was further subdivided into an *on–in–under* group ($n = 10$ per age group) and a *top–middle–bottom* group ($n = 12$ per age group). In each age by word set group, half the children were male and half were female.

2.1.2. Design

There were two between-subject factors, Age (3;8 and 4;1 years) and Condition (baseline and language), and one within-subject factor, Location (top, middle, or bottom). The language condition was comprised of two groups, one hearing the word set *on*, *in*, *under* (OIU) and one hearing the word set *top*, *middle*, *bottom* (TMB). The primary dependent measure was the proportion of correct responses on the six search trials. Children were tested twice at each location.

2.1.3. Materials

There were two boxes, a Hiding box and a Finding box, placed about 2 ft apart on the floor (Fig. 1). Each box was roughly 15 in. high, 12 in. wide, and 7 in. deep. Four cards were created for each box, each of which was a clear acrylic 5 in. × 7 in. picture frame with construction paper inside. The Hiding box was white with light gray cards; the Finding box was light blue with dark blue cards. One of the four cards for each box had a star on the back, making it the "winner." At all times there was a card placed on, in, and under each box, only one of which was the winner.

2.1.4. Procedure

2.1.4.1. Orientation. Children were first shown the Hiding box and the four accompanying cards, front and back. Three ordinary cards were placed on, in, and under the Hiding box, and a fourth card with a star on its back (the "winner") was placed in front of the box. In the Baseline condition, the experimenter said for each card: "Let's look at this one. Does it have anything on the back?" In the Language condition, the experimenter used the specific spatial term for each location, for example, "Let's look at the one *on* (or *in* or *under*) the box. Does it have anything on the back?" (For the TMB word set, the phrasings were *at the top of the box*, *at the middle of the box*, and *at the bottom of the box*.) This was repeated with the Finding box and its cards.

The experimenter then explained the mapping task. Children were told that the winners were always put in the same place at the two boxes and given a practice trial. The experimenter placed the winner to the right of the Hiding box, saying either "I'm putting this winner right here" (baseline condition) or "I'm putting this winner *next to* the box" (for the OIU word set) or "I'm putting this winner *at the side of* the box" (for the TMB word set). The experimenter then put the winner at the Finding box in the corresponding place, saying: "And this winner goes right here, in the very same place." The Finding box instruction was the same for all children; the experimenter never applied spatial relational terms to the Finding box during the mapping task. Children were asked to find the winner at the Finding box, and to retrieve the original winner at the Hiding box. This practice trial was used to provide a clear demonstration of how to play the game and what children were expected to do.

2.1.4.2. Search trials. As children watched, the experimenter placed the winner on, in, or under the Hiding box, saying either "I'm putting the winner right here" (baseline condition) or "I'm putting the winner [*on*/*in*/*under* or *at the top of at the middle of* *at the bottom of*] the box" (for the two language groups). The experimenter then asked children to close their eyes while hiding the other winner at the Finding box. Children opened their eyes and searched for the winner "in the very same place" at the Finding box. Children were tested at each location once before being tested a second time for a particular location. Two orderings of placements were used. Between the fourth and fifth trials there was a catch trial in which the winners were placed next to the boxes, just as in the practice trial. Children performed well (80% correct) on this trial, confirming that they were attentive and understood the search task.

2.1.4.3. Retrieval trials. After children found the winner at the finding box, they retrieved the winner from the Hiding box (i.e., the one they had seen being placed). This is a standard procedure for mapping tasks, and is typically used to assess children's memory for the initial location (e.g., DeLoache, 1987).

2.1.4.4. Open-ended question. After completing the mapping task, children were asked: "How did you know where to look? How did you know which card would be the winner?"

2.1.4.5. Scoring. Children were scored correct on a search or retrieval trial only if they searched at the correct location first, although they were allowed to look for the winners until they found them. The main dependent measure was the proportion correct across the six search trials. We also performed two comparisons to chance performance ($p = .33$): (1) whether age $\times$ language groups performed reliably above chance, and (2) the number of individual children in each group who performed above chance (i.e., a minimum of 5 out of 6 trials correct, as dictated by the binomial distribution with $p = .33$). There were no effects of gender, nor of order of presentation, nor was there any change in search or retrieval performance across trials in this study or in any of the subsequent studies.

2.1.4.6. Scoring children's comments. Children's responses to the final open-ended question (how they knew where to look) were analyzed in two ways. First, the comments were given to 11 raters who rated them on a scale from 1 (not informative) to 7 (informative) (Cronbach's $\alpha = .95$). The raters were made familiar with the mapping task and were given written versions of the children's responses and descriptions of their accompanying gestures. They had no access to further information about the children, such as their age, condition, or task performance. The second analysis was of the content of children's comments. These were analyzed for references to spatial information (e.g., the use of words such as *here* or *there*, or gestures to a location), and for references to similarity between the boxes (the word *same*, or gestures to matching locations in the two boxes). In addition, to check for possible surface strategies, we also analyzed children's comments for any reference to the experimenter's language—either the use of the specific terms used in the language condition (*on*, *in*, *under*, *top*, *middle*, and *bottom*) or a direct reference to what the experimenter said.

2.2. Results

2.2.1. Search trials

The key question was whether hearing spatial relational language would facilitate children's mapping performance. The performance of the 3;8-year-olds showed a clear language advantage: the language children were 72% correct, but the baseline group was only 45% correct. A 2 (Age) × 2 (Language) × 3 (Location) mixed model ANOVA showed a main effect of language, $F(1,84) = 9.53$, $p < .005$, $MSE = .20$ (see Fig. 2 and Table 2). There was no age effect, nor a language by age interaction, $F(1,84) = 2.94$, $p = .09$. Planned contrasts of condition within each age level confirmed the language effect among 3;8-year-olds, $F(1,84) = 11.53$, $p < .005$. Among 4;1-year-olds, the language (71%) and baseline (64%) conditions did not differ significantly. All four groups performed reliably above chance levels (33%), including the younger baseline group, $t(21) = 2.35$, $p < .05$. Analyses over individual children showed that 61% of the language children performed above chance (33%)—that is, they searched correctly on at least 5 out of 6 trials—as compared with 20% of the baseline children, $\chi^2(1, n = 88) = 15.23$, $p < .001$.

Further ANOVAs confirmed that same pattern held for the two word sets considered separately. Language children performed better than baseline children in both the TMB group and the OIU group, $F(1,82) = 7.71$, $p < .05$, $MSE = .07$, and $F(1,82) = 4.83$, $p < .01$, respectively.

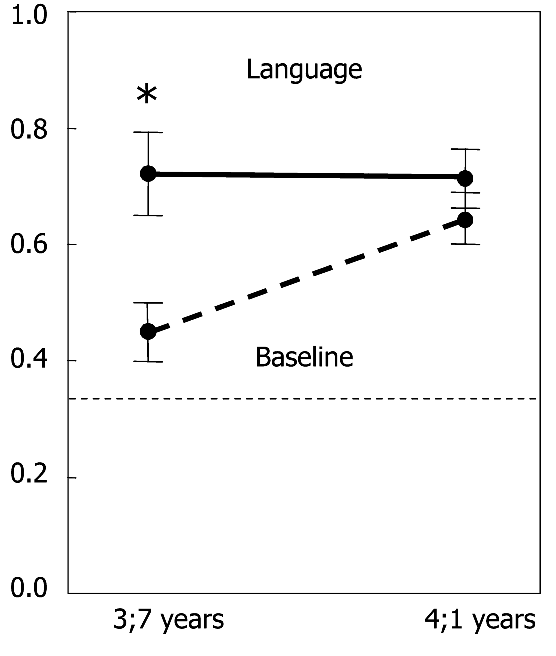

Fig. 2. Children's proportion correct on the search trials for Experiment 1.

Table 2
Performance on search and retrieval trials for language and baseline conditions by age across experiments: mean (SD)

Experiment	Age (years; months)	Search trials			Retrieval trials		
		Baseline	On–In–Under	Top–Middle–Bottom	Baseline	On–In–Under	Top–Middle–Bottom
1. Neutral	3;8	.45 (.24)[a]	.67 (.34)[a,*]	.76 (.33)[a,*]	.65 (.22)[a]	.92 (.12)[a,*]	.88 (.23)[a,*]
	4;1	.64 (.19)[a]	.73 (.27)[a]	.69 (.24)[a]	.86 (.19)[a]	.92 (12)[a]	.81 (.26)[a]
2. Neutral Pre	3;8	.48 (.20)[a]	.67 (.16)[a]		.64 (.22)[a]	.76 (.20)[a]	
	4;3	.60 (.22)[a]	.76 (.28)[a]		.78 (.24)[a]	.84 (.29)[a]	
3. X-Map	4;1	.28 (.19)	.32 (.20)		.53 (.22)[a]	.67 (.24)[a]	
	4;7	.45 (.22)	.65 (.23)[a,*]		.82 (.21)[a]	.85 (.15)[a]	
	5;2	.53 (.29)	.73 (.18)[a,*]		.73 (.32)[a]	.90 (.22)[a]	
4. X-Map	3;7		.44 (.19)	.57 (.19)[a]		.57 (.21)[a]	.65 (.22)[a]
	4;2		.35 (.18)	.67 (.27)[a,*]		.51 (.29)	.76 (.29)[a,*]
5. Neutral Pre	3;7	.50 (.22)[a]		.60 (.26)[a]	.53 (.17)[a]		.67 (.28)[a]
	4;3	.57 (.22)[a]		.80 (.15)[a,*]	.72 (.18)[a]		.88 (.14)[a]
5. Transfer	3;7	.45 (.24)		.70 (.27)[a,*]	.52 (.24)[a]		.72 (.21)[a]
	4;3	.67 (.24)[a]		.68 (.18)[a]	.75 (.24)[a]		.83 (.21)[a]
5. Retention	3;7	.45 (.14)[a]		.70 (.27)[a,*]	.50 (.21)[a]		.67 (.21)[a]
	4;3	.65 (.27)[a]		82 (.21)[a]	.82 (.24)[a]		.90 (.18)[a]

Note. Experiment 1 used a neutral-objects mapping task. Experiments 2 and 5 used a neutral-objects mapping task with language provided prior to (rather than during) the task. Experiments 3 and 4 used cross-mapped tasks. Experiment 5 also used immediate transfer and delayed retention neutral-objects tasks.

[a] Performance better than chance (.33).

[*] $p < .05$ for the contrast with baseline.

There was also a main effect of location, $F(2,168) = 3.53, p < .05, MSE = .01$. Children showed a bias towards the top location, both when it was correct ($M_{\text{top}} = 70\%$, $M_{\text{middle}} = 62\%$, $M_{\text{bottom}} = 57\%$), and when it was not (63% of errors when middle was correct, 70% of errors when bottom was correct, $p < .05$, $p < .005$, respectively).[1] (However, because all locations were correct equally often, a bias towards the top location results in the same overall score as choosing randomly (i.e., 33%).)

2.2.2. Retrieval trials

The 3;8-year-olds also showed a language advantage when retrieving the original winner they had seen being placed (89% vs. 65% correct for the language vs. baseline groups), as shown in Table 2. A 2 (Age) × 2 (Language) × 3 (Location) ANOVA confirmed a main effect of language, $F(1,84) = 7.88$, $p < .01$, $MSE = .12$. The age effect was non-significant, $F(1,84) = 3.72$, $p = .06$, although there was a significant interaction between the two, $F(1,84) = 7.88$, $p < .01$. Planned contrasts within each age

[1] Here, as in all the error analyses, a binomial test was used to test against a chance level of 50% (because there are two possible errors).

group confirmed the language advantage among 3;8-year-olds, $F(1,84) = 15.75$, $p < .001$. The 4;1-year-olds performed equally well across the language (86%) and baseline (86%) conditions. All groups performed above chance, minimum $t(21) = 6.86$, $p < .001$. However, as with the search trials, individual analyses revealed that more language (84%) than baseline (52%) children performed above chance (33%), $\chi^2(1, n = 88) = 10.27$, $p < .005$. As with the search trials, both word sets conferred an advantage relative to the baseline condition.

2.2.3. Children's comments

Half (53%) the children responded to the final question "How did you know where to look?" The group that responded was much like the group that did not in their condition, gender, and distribution of scores. There was a considerable range of specificity in children's answers, from "cause I just did" to "I watched where you put in the first

Table 3
Comments from Experiment 1, grouped by content and condition

Content	*n*	*M*	Language	*n*	*M*	Baseline
Similarity & Space	8	.92	It's in the same place (3) I watched where you put in the first one, and then you just put it in the same place (3) I found them there and there, there and there, there and there. They're the same 'cause they match each other (gesturing) (1) (pointed to top of both boxes) I'm tricky because I just know (1)	3	.67	In the same place (1) (pointed to top of both boxes) Same place (1) You showed me over there and I found in the other one (1)
Space only	6	.75	Because I always think where it is (2) (pointed to top of one box) Here (1) I knew those boxes were there sort of hiding. I looked in back for the star (1) I can see where you put it. I just knew (1) You told me where it was (1)*	4	.63	Because I thought it was there (1) (pointed to top, middle of one box) (1) Because I knew all the places over there [hiding box] (1) Because you tell me... I have a good sense of space (1)*
Other	10	.62	I don't know/guessed (3) Because I know/ I'm smart (4) An one an better all one (1) I learned it. You taught me, you helped me play that game (1) Because I heard (1)*	16	.53	I don't know/guessed (8) Because I know/ I'm smart (5) My dad told me. (2) Learned... wanted to each time find it (1)
Total	24	.75		23	.57	

n, number of children making this kind of comment (each child appears only once in the table).
M, mean performance on the search task for children who made this kind of comment.
* Comments referring to the experimenter's language per se.

one, and then you just put it in the same place.'' Table 3 presents the children's comments, categorized according to whether they referred to spatial similarity, to spatial information only, or to neither (no child mentioned similarity other than to refer to spatial similarity). Each child appears only once in the table and highly similar statements are grouped. No child repeated the specific terms used in the language condition. However, three children (two in the language condition) repeated or referred directly to what the experimenter had said (e.g., ''you told me where it was''), and these comments could conceivably reflect a surface cuing approach to the language. With these comments excluded, children who referred to spatial information performed better on the search trials than those who did not (75% vs. 55% correct), $t(42) = 2.46$, $p < .05$. Language children (59%) were more likely than baseline children (27%) to mention spatial information, $\chi^2(1, n = 44) = 4.54$, $p < .05$.

The ratings of children's comments showed that those that referenced both spatial location and similarity of location ($M = 5.3$) were rated as more informative than comments that referred only to space ($M = 2.8$), which in turn were considered more informative than comments referring to neither similarity nor space ($M = 1.8$), all Bonferroni-adjusted p's $< .01$. The ratings showed a positive correlation with children's performance on the search task $r(47) = .36$, $p < .05$.

2.3. Discussion

The 3;8-year-old children who heard either set of spatial language terms performed better on a spatial relational mapping task than their age mates who did not. These young children were also better than those in the baseline condition at retrieving the original winner in the first box. Children's comments also reflect an influence of language on their conceptual understanding. Children hearing spatial terms made more comments indicative of spatial insight and/or of having noticed the mapping between the boxes than did children in the baseline condition. Overall, the results are consistent with the claim that relational language invites a more delineated representation of the spatial relations—a representation that permitted a specific analogical mapping between the two boxes.

2.3.1. Alternative accounts

We are suggesting that the effect of hearing spatial terms came about because children were guided by the meanings of the spatial terms in interpreting the scenes. However, three compatible but weaker explanations might be sufficient to explain the effects without needing to appeal to relational knowledge. First, perhaps there was a ''task engagement'' effect whereby hearing the language labels simply made the task more interesting. A second possibility is that language added a unique attribute that served as a distinctive cue to the proper correspondence—roughly comparable to adding a matching whistle or tone. The words were not applied to the Finding Box during the task, but because they were applied to both boxes during the introduction this ''surface cuing'' account cannot be disregarded. A third alternative is that children were using the meanings of the words to remove the need to map. On this ''verbal bypass'' account, language children simply listened to the word used and then looked in that place at the finding box. Thus, whereas the baseline children were faced with a challenging mapping task, the language children could carry out a simple language-comprehension task.

Some aspects of the current results argue against the alternative accounts. Children performed well on the catch trial across all ages and conditions, suggesting that task engagement was not the problem. Against the surface cuing account, no child mentioned the specific words used by the experimenter, and only a few even mentioned the experimenter's use of language. Even with these children omitted, the language group referred to spatial locations or spatial similarity between the boxes more often than the baseline children. This suggests: (1) that the spatial labels operated at a conceptual level, contrary to the surface cuing account; and (2) that the labels served to inform the mapping between the boxes, not to bypass the mapping process, contrary to the verbal bypass account. Had the language children simply been treating the labels simply as distinctive features (analogous to whistles and bells), they might have said "Because you made the same sound," or "Because you said the same thing;" but for children to note that "they were in the same *place*" requires them to have processed the fact that the terms refer to locations—that is, to have invoked the meanings of the terms. Overall, we believe the results fit better with the semantic account than with any of the three non-semantic explanations. We present further evidence on this issue in the next study.

Finally, we note that our perspective is compatible with one that emphasizes the pragmatic and attentional aspects of using the spatial terms. Hearing spatial language may have guided children's attention towards spatial information generally, over and above the specific spatial relations conveyed by the terms. As Gelman and Greeno (1989) have noted, young children often lack full interpretive competence—they may underperform because they fail to understand the task (see also Tomasello, 1999; Zelazo & Frye, 1998). The linguistic context can highlight particular dimensions and hence lead children to invoke some strategies and not others.

3. Experiment 2

If the effects of spatial language in Experiment 1 occurred on the semantic-conceptual level (as opposed to via one of the non-semantic routes discussed earlier) then it should be possible to improve children's performance by giving them spatial language *before* the mapping task. In Experiment 2, prior to engaging in the mapping task, children in the language group placed toys at one box as directed by spatial language (e.g., "can you put this *on* the box?"). The spatial terms were applied only to the Hiding box. (Language children did not see the Finding box until the mapping task, during which none of the spatial language terms were used.) The baseline children were shown both the Hiding and the Finding boxes during this training phase, but received only general language and gesture (e.g., "can you put this one right here?" [pointing]) rather than receiving specific spatial terms. After training, all children were given the mapping task from Experiment 1 using the baseline procedure—no spatial relational terms were used during the mapping task itself.

If the effect of spatial language in Experiment 1 is to invite a relational encoding, as we suggest, then children who hear spatial language prior to the mapping task should form better spatial relational representations and should therefore perform better than those who do not. However, if the spatial language advantage in Experiment 1 was due to surface cue matching, then there will be no advantage for the language group, because the spatial terms will never be applied to both boxes. Also, because spatial language will not be used during the mapping task, if the language effect in Experiment 1 resulted from children bypassing the mapping and simply attending to the experimenter's labels, then the effect should disappear here. Likewise, if the language effect arose because the spatial terms added interest to the task, then it should not appear here.

3.1. Method

3.1.1. Participants

The participants were 48 children from the same population as in Experiment 1. The younger group averaged 3;8 years old (range: 3;5–3;11 years), and the older group of children averaged 4;3 years old (range: 4;0–4;7 years). Half were female and half were male. Children were randomly assigned to either an *on*, *in*, *under* language group or a baseline group. There were 11 children in the younger baseline group, 12 in the younger language group, 13 in the older baseline group, and 12 in the older language group.

3.1.2. Design

Age (3;8 and 4;3), and Condition (language and baseline) were between-subject factors, and Location (top, middle or bottom) was a within-subjects factor. The primary dependent measure was the proportion of correct responses on the six search trials (two trials for each location).

3.1.3. Materials

This study used the same boxes as in Experiment 1, but instead of cards, we used toy animals (pigs and chickens) that could be opened. To designate a winner, we placed a small three-dimensional plastic star inside one toy animal.

3.1.4. Procedure

3.1.4.1. Pre-task training. Children were given an introductory task that varied with condition. Language children were shown only the Hiding box (empty), and the six plastic animals. The experimenter handed children one of the toys and asked them to place it at the box: "Can you put this *on* [/*in*/*under*] the box?" After children placed a toy, the experimenter removed it and began the next trial. They were given three trials for each location at the box (nine trials in total), and all three locations were visited before a location was repeated. Children were not corrected, but were nearly always accurate. The Baseline children saw both the Hiding and Finding boxes and the six animals. The children played a placement game as in the language condition, except that instead of naming the location, the experimenter pointed to where the children should put the animal and said: "Can you put this one right here?" Children were given three trials at the Finding box, then three with the Hiding box, then three more with the Finding box (i.e., nine trials in total). After completing one of the two tasks, the experimenter set up the boxes for the mapping task.

3.1.4.2. Mapping task. The mapping task was the same for both conditions, and highly similar to Experiment 1's baseline condition. Three plastic chickens were placed at the Hiding box and three pigs were placed at the Finding box. The experimenter hid a plastic star in one of the chickens at the Hiding box and said: "I'm making this one the winner, and putting it right here." Then, after closing their eyes for the hiding event, children were asked to "find the winner in the same place over here" at the Finding box. As before, after children found the winner at the Finding box, they retrieved the original winner from the Hiding box. Also as before, children were given six search and retrieval trial pairs, two per location, and a catch trial after the fourth trail (on which all children performed well: 95% correct).

3.2. Results

3.2.1. Search trials

Children in the language condition (72% correct) performed better than those in the baseline condition (54%), even though the spatial terms were presented only initially, and not during the task itself. A 2 (Age) × 2 (Language) × 3 (Location) ANOVA confirmed the predicted effect of language, $F(1,44) = 7.23$, $p < .05$, $MSE = .15$ (see Fig. 3 and Table 2). There was no effect of age, nor any interaction with age. Planned contrasts within each age level showed no significant language effects at 3;8 (67% correct for language, 48% for baseline), or at 4;3 (76% vs. 60%), $F(1,44) = 3.90$, $p = .05$, $F(1,44) = 3.33$, $p = .07$, respectively. As in the first study, all groups performed above chance (33%) levels, including the younger baseline group, $t(10) = 2.51$, $p < .05$. Over twice as many language children (50%) performed above chance (33%) as did baseline children (21%), $\chi^2(n = 44) = 4.46$, $p < .05$. There was also a main effect of location, $F(1,44) = 6.17$, $p < .005$, $MSE = .08$, with children more often correct at the topmost location (75%) than at either the middle location (57%) or the bottom location (57%), $p < .05$.

3.2.2. Retrieval trials

A 2 (Age) × 2 (Language) × 3 (Location) ANOVA examining children's retrieval trial performance revealed no effect of language or age nor any interaction between them. Planned contrasts showed no differences due to language at 3;8 (76% vs. 64% correct for the language vs. baseline groups), nor at 4;3 (84% vs. 78%). All groups performed reliably above chance, all t's > 4.6, p's < .005. Comparisons of individual performance against chance revealed no significant difference in performance by language (66%) and baseline (46%) children. There was no significant effect of location, $F(1,44) = 3.10$, $p = .05$, $MSE = .06$, with children tending to be most correct at the top (83%) followed by the middle (73%) and then the bottom (71%).

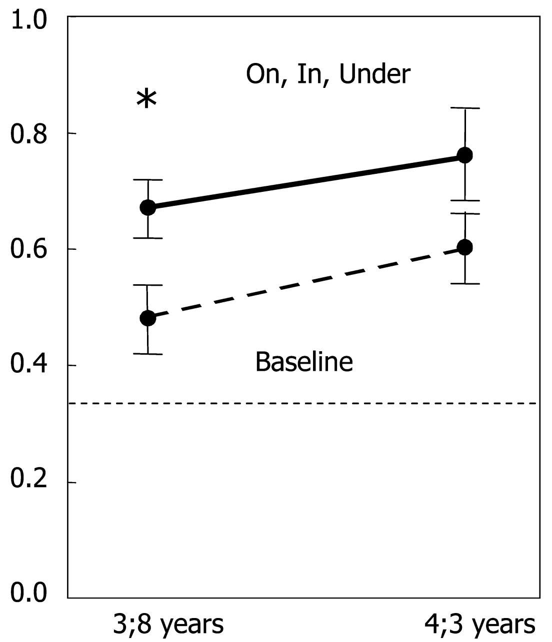

Fig. 3. Children's proportion correct on the search trials for Experiment 2.

3.3. Discussion

Even modest exposure to spatial language can yield an advantage for children's mapping performance. Overall, language children performed better than baseline children on the mapping task, despite only hearing spatial language during the brief introductory task at the beginning of the session, and not during the actual mapping task.

These results provide evidence that the language effect occurs at the conceptual level, and is not merely due to one of the surface strategies discussed above. Because the language children received labels only for the Hiding box, it is unlikely that their superior performance could have resulted from matching surface cues between the boxes. Further, because the two groups were treated identically during the actual mapping task, we can discount the "task engagement" explanation for the language benefit. Finally, we can rule out the "verbal bypass" account, because the spatial terms were not used during the mapping trials. Overall, we think it most plausible that the effect of spatial terms is to invite a conceptual representation of the spatial structure of the Hiding box, which then serves as a basis for a relational mapping between the two boxes.

4. Experiment 3

Experiment 3 tested the second of our three predictions, that effects of spatial relational language will be seen at later ages for more difficult tasks. We made the spatial mapping task more challenging to examine potential advantages of spatial language for older children. We predicted that older children would benefit from the explicit use of spatial relational language on a more challenging task just as young children did on the simple task described above.

To generate a more difficult spatial mapping task, we used Gentner and Toupin's (1986) cross-mapping methodology to introduce a mismatch between object matches and relational correspondences. An example of a cross-mapping is the pair OXO and XMX. The X in the first string has a clear object match to the X's in the second string, and a relational match to the M (because both are in the middle position). Such cross-mappings between object matches and relational matches are challenging because the salient object match (e.g,. X to X) competes with the correspondences dictated by the relational match (X $\rightarrow$ M and O $\rightarrow$ X; Gentner & Toupin, 1986; Markman & Gentner, 1993; Ross, 1987). Cross-mappings are particularly difficult for children, who are more likely than adults to focus on object properties (Gentner, 1988; Halford, 1993; Rattermann & Gentner, 1998a, 1998b). For example, Gentner and Toupin (1986) asked 5-year-olds to act out a story with three characters, and then asked them to reenact the story using new characters. When the characters were similar to and played the same roles as the original characters (making a literal similarity match) the children were highly accurate. In contrast, when the characters in the story were similar to the original characters, but played different roles (a cross-mapping), children performed poorly. Other developmental results have converged on this finding that cross-mappings are difficult for children (Gentner & Rattermann, 1991). Accordingly, the following experiment presented children with a cross-mapped version of the box task, which should be more difficult than the prior neutral objects task. The prediction is that spatial language will facilitate mapping performance for older children.

Experiment 3 was in many respects similar to Experiment 1. Children engaged in the hiding and finding task using the two boxes, and were either in a Language (*on*, *in*, *under*) or Baseline condition. It differed in that 4-, $4\frac{1}{2}$-, and 5-year-old children were tested, and the cards placed at the boxes had pictures on them, with identical pictures in different relative positions so that there was a cross-mapping (see Fig. 1). If our hypothesis that spatial relational language invites relational encodings is correct, then we should see an advantage for the language group over the baseline group. Also, based on prior findings of a relational shift in children's perception of similarity (Gentner & Rattermann, 1991), we predict poorer baseline performance among the 4-year-olds here than in Experiment 1, because the cross-mapped objects should pose a significant challenge.

4.1. Method

4.1.1. Participant

Participants were 60 children from the same population as in Experiment 1, in three age groups of 20 children each: 4;1 years old (range: 4;0–4;3); 4;7 years old (range: 4;5–4;9); and 5;2 years old (range: 4;11–5;4). Half the children were randomly assigned to the Language group, and half to the Baseline group. Equal numbers of children within each age and condition group were male and female.

4.1.2. Design

There were two between-subjects factors, Age (4;1, 4;7, and 5;2) and Condition (language and baseline), and one within-subject factor, Location (top, middle, and bottom). The main dependent measure was the proportion of correct responses on the six search trials (two for each location).

4.1.3. Materials and procedure

The boxes were those used in the Experiment 1. However, instead of identical solid-colored cards, three identical pairs of brightly colored pictures were created. The pictures were placed in the boxes such that matching pictures were in mismatched locations (Fig. 1). For instance, a card with a picture of fish might be placed in the middle position at one box and in the bottom position at the second box. The "winner" was designated by temporarily affixing a star to the back of one of the cards.

The procedure was as in Experiment 1. Children were given an orientation phase, followed by six search trials and the accompanying retrieval trials. There was an easy catch trial between the fourth and fifth trials on which children performed well (78%). The language group heard the word set *on*, *in*, and *under* (e.g., "I'm putting the winner *under* the box), and the baseline group heard "I'm putting the winner right here."

4.2. Results

4.2.1. Search trials

As expected, the cross-mapped task was considerably more difficult than the neutral objects task of Experiment 1. The key prediction was borne out: we found a language advantage among older children (see Fig. 4 and Table 2). A 3 (Age) × 2 (Language) × 3 (Location) ANOVA showed a main effect of language, $F(1,54) = 6.34$, $p < .05$, $MSE = .15$, and a main effect of age, $F(2,54) = 12.19$, $p < .001$. Their interaction was not significant. Planned contrasts of language effects within each age level showed no difference at 4;1 (32% correct for the language group, 28% for the baseline group). However, there were significant language effects at 4;7 (65% vs. 45%), $F(1,54) = 4.05$, $p < .05$, and at 5;2 (73% vs. 53%), $F(1,54) = 4.05$, $p < .05$. The performance of both groups of 4;1-year-olds did not differ from chance (33%) (t 's < 1), nor did that of the 4;7-year-old baseline group, $t(9) = 1.66$, $p = .13$, nor the 5;2-year-old baseline group, $t(9) = 2.17$, $p = .06$. Only the 4;7- and 5;2-year-old language groups performed above chance, minimum $t(9) = 4.39$, $p < .005$.

There was a main effect of location, $F(2,108) = 4.26$, $p < .05$, $MSE = 0.12$, which was moderated by an interaction with Age $F(4,108) = 2.61$, $p < .05$. Whereas the older two groups showed no location bias, the 4;1-year-olds tended to choose the topmost location ($M_{\text{top}} = 53\%$, $M_{\text{middle}} = 18\%$, $M_{\text{bottom}} = 20\%$).

We also examined children's object errors: that is, errors due to choosing the matching object instead of the corresponding relative location. (These errors are independent of location errors: children were given two trials per location, and an

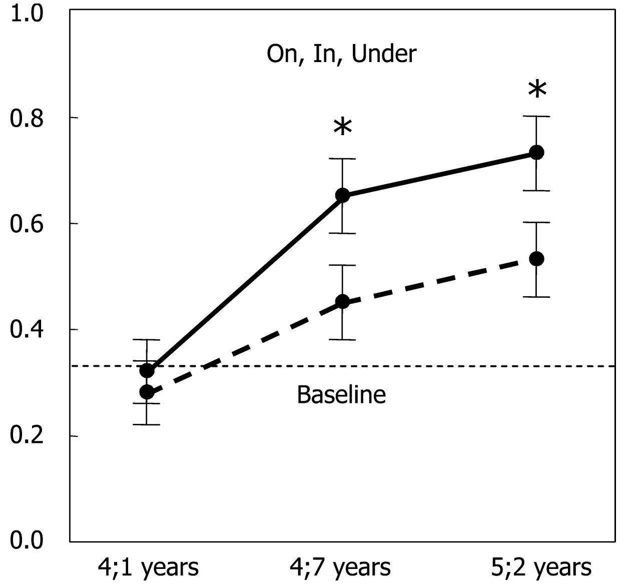

Fig. 4. Children's proportion correct on the search trials for Experiment 3.

object error could be made on any trial.) The 5;2-year-olds often made object errors (64% of their errors, respectively), $p < .05$, by binomial tests against chance (50%, see Footnote 1; 61% of the 4;7-year-olds errors were object errors, $p = .07$). The 4;1-year-olds reliably made object errors on their first trial (9/13 errors, or 69%, $p < .05$), but not thereafter (31/71 errors, or 44%). These children, upon being shown that the object match was incorrect, appeared to abandon the mapping effort (they performed at chance levels on the search trials). Many appeared to switch to a fixed strategy such as sequentially searching in a top, middle, bottom order on each trial. Such fixed search strategies partly account for the 4;1-year-olds' high number of errors to the top location, and account for all of their catch trial errors.

4.2.2. Retrieval trials

Older children performed better on the retrieval trials than younger children. A 3 (Age) × 2 (Language) × 3 (Location) ANOVA showed an effect of age, $F(2,54) = 6.32$, $p < .005$, $MSE = .16$, but no significant effect of language, $F(1,54) = 3.46$, $p = .07$ (see Table 2). Their interaction was not reliable, nor were planned contrasts of language effects within each age level. All groups performed at greater than chance (33%) levels, including the youngest baseline group, $t(9) = 2.88$, $p < .05$. At the individual level, 70% of the language children and 47% of the baseline children performed above chance, a non-significant difference, $\chi^2(1, n = 60) = 3.36$, $p = .07$. There was no effect of location.

Children's retrievals were also examined for reverse object errors: that is, choosing the card at the Hiding box whose picture matched that of the winner at the Finding box. Because children were always allowed to search until they found the winner at the Finding box, they could approach the retrieval trials by mapping *back* from the Finding box to the Hiding box. If they adopted this strategy and mapped on the basis of object similarity, the result would be a reverse object error. Among the youngest children, 65% of the errors were reverse object errors (32 out of 49 errors, $p < .05$ by a binomial with 50% as chance). The 4;7-year-olds (45%) and 5;2-year-olds (41%) did not systematically make such errors.

4.3. Discussion

As in the first two studies, children who heard spatial relational language performed better on a spatial mapping task. It is clear that the cross-mapped task with its competing object matches was more difficult than the neutral objects mapping task (from Experiments 1 and 2), consistent with prior findings of a relational shift in children's processing (Gentner, 1988; Gentner & Rattermann, 1991; Halford, 1993). When given neutral objects in the earlier experiments, 4;1-year-old children performed well above chance even in the baseline condition. In contrast, the 4;1-year-olds in the current study performed at chance levels regardless of condition. However, among 4;7- and 5;2-year-olds, the greater difficulty of the cross-mapped task was compensated for by the presence of spatial language. This bears out our second prediction—that when the mapping task is made more difficult, older children will show a benefit from overt relational language.

The 4;1-year-olds' performance in this difficult task is interesting. As expected, they did not appear to notice relational similarities across the boxes; but more surprisingly, they did not consistently search based on object similarity either. They reliably made object errors on their first trial but then appeared to abandon the mapping task, often resorting to searching in a fixed pattern. However, they often relied on object similarity in their *retrieval trial* performance. Having seen the winner at the Finding box, they reliably made "reverse object errors," in which they chose the card in the Hiding box that looked the same as the winner in the Finding box.[2] The overall pattern of the 4;1-year-olds' performance suggests that they were not able to maintain a consistent relational mapping.

The 4;1-year-olds' poor performance in the mapping task (even with the support of *on*, *in* and *under*) could indicate a developmental limitation. For example, perhaps young children lack the processing capacity to carry out a relational mapping when there are competing object matches; or perhaps they lack the ability to inhibit object matches, possibly due to incomplete development of prefrontal circuitry. A third possibility, explored in Experiment 4 and in simulation studies, is that the children's relational representations were insufficiently elaborated to prevail against a competing object match.

---

[2] The high number of reverse object errors in this study sounds a cautionary note concerning the interpretation of retrieval trials in children's search tasks. The standard assumption is that they serve as tests of children's memory for the original item's location; clearly this assumption would not hold in the current study, and quite possibly for others.

5. Experiment 4

Experiment 4 tested our third prediction—that the semantics of the terms should determine what effects language might have on children's mapping performance. Specifically, we contrasted the performance of children hearing *on*, *in*, *under* with children hearing *top*, *middle*, *bottom* on the cross-mapped task from Experiment 3. These two word sets have very different properties. *Top*, *middle*, *bottom* conveys an integrated system set of relations. Each term refers to one relation within a larger vertically oriented structure, and the terms are monotonically ordered within that larger structure. For example, the American Heritage Dictionary of the English Language (2000) defines *top* as "the uppermost part, point, surface, or end," *bottom* as "lowest in position, rank, or place," *middle* as "equally distant from extremes or limits; central." In contrast, the terms *on*, *in*, *under* each convey a separate first-order relation between a figure and a ground: support from below (*on*), containment (*in*), and coverage from above (*under*) (e.g., Herskovits, 1986). Thus *on*, *in*, and *under* serve as three separate spatial relations, or at best as a pair (*on*, *under*) with a separate first-order relation (*in*) interpolated. This is not to deny that their meanings interact—for example, the boundary between *on* and *in* has been the subject of cross-linguistic study (Bowerman & Pederson, 1992; Coventry, 2001; Feist, 2000). However, the *on–in* continuum concerns support vs. containment, not relative spatial position. The three terms do not form a connected spatial system, as *top*, *middle*, *bottom* do. Prior studies have borne out the prediction of structure-mapping theory that connected systems of relations are implicitly favored over independent relations in analogical processing (Clement & Gentner, 1991) and are more likely to prevail against cross-mapped object matches (Gentner & Rattermann, 1991; Markman & Gentner, 1993). We therefore predicted that children hearing *top*, *middle*, *bottom* would perform better than those hearing *on*, *in*, *under* on the cross-mapped mapping task from Experiment 3.

5.1. Method

5.1.1. Participants

Participants were 48 children from the same population as in Experiment 1. There were two age groups: 3;7 years old (range: 3;4–3;9) and 4;2 years old (range: 4;0–4;5). Half the children heard *top*, *middle*, *bottom* (TMB), and half heard *on*, *in*, *under* (OIU), making 12 children in each Age by Word set group. Half of each group of 12 were male and half were female.

5.1.2. Design

The design was 2 (Age: 3;7 and 4;2) × 2 (Word set: OIU and TMB) × 3 (Location). Age and word set varied between-subjects and location varied within-subjects.

5.1.3. Materials and procedure

The method was as in Experiment 3, except that the contrast here was between two different language conditions, rather than between language and baseline. The materials were identical. There was one difference in procedure, namely the addition of an introduction laying out the set of terms for that child. The experimenter said, for example: "Now we're going to play the top–middle–bottom game! Have you ever played the top–middle–bottom game?" The experimenter then showed the children the boxes, and pointed out their locations, one box at a time: "See, this one has a top, and a middle, and a bottom. And this one has a top, a middle, and a bottom. That's why I call it the top–middle–bottom game." A similar procedure was followed for the *on–in–under* group. The experimenter then began the orientation phase of the experiment and continued exactly as in Experiment 3. During the remainder of the study, both groups heard the words individually and applied to the Hiding Box: for example, "I'm putting the winner *at the top of* the box."

5.2. Results

5.2.1. Search trials

The results were straightforward: children who heard *top*, *middle*, and *bottom* (TMB; 62% correct) performed substantially better than children who heard *on*, *in*, and *under* (OIU; 40%) (see Fig. 5 and Table 2). A 2 (Age) × 2 (Word set) × 3 (Location) ANOVA confirmed the advantage of TMB over OIU, $F(1,44) = 13.31$, $p < .005$, $MSE = .13$. There was no effect of, nor interaction with, age. Planned contrasts of word set effects within each age level showed a significant TMB advantage at 4;2 (TMB: 67%; OIU: 35%), $F(1,44) = 13.76$, $p < .001$, $MSE = .09$, but not at 3;7 (TMB: 57%; OIU: 44%). Both the 3;7- and 4;2-year-old TMB groups performed reliably above chance levels, $t(11) = 4.21$, $p < .005$ and $t(11) = 4.34$, $p < .005$,

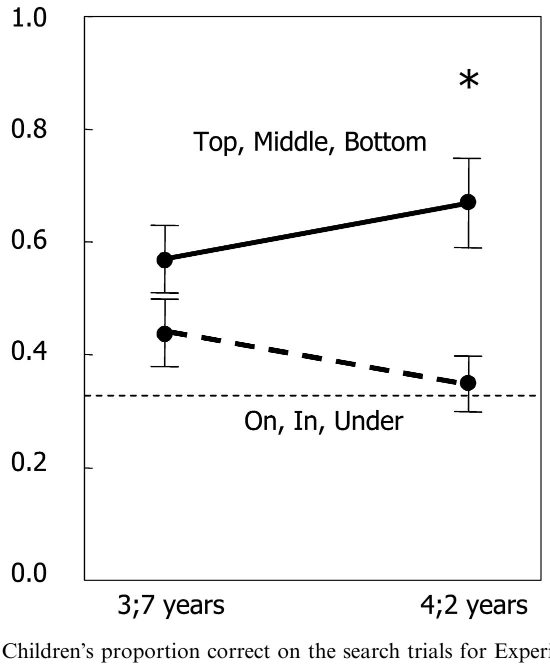

Fig. 5. Children's proportion correct on the search trials for Experiment 4.

respectively. Within the OIU group, neither age group performed better than chance. Finally, 7 of the TMB children met the 5/6 criterion for above-chance performance, but none of the OIU children did so, $p < .01$ by Fisher's exact test.

There was no main effect of location, but there was a word set by location interaction, $F(2, 88) = 8.24$, $p < 005$, $MSE = .12$. Children hearing *on*, *in*, and *under* were most correct at the top location (58, 33, and 27%, respectively), while children hearing *top*, *middle*, and *bottom* were most correct at the bottom location (52, 54, and 79%, respectively). Children reliably made object errors: 83 of the 143 errors (58%) were object errors (above chance (50%) by a binomial, $p < .05$).

5.2.2. Retrieval trials

The TMB group also showed an advantage relative to the OIU group (71% vs. 54%) in retrieving the original winner in the Hiding box. A 2 (Age) × 2 (Word set) × 3 (Location) ANOVA showed a main effect of word set, $F(1, 44) = 5.19$, $p < .05$, $MSE = .19$. There was no main effect of, nor interaction with, age. Planned contrasts of word set effects within each age level showed a significant difference at 4;2 (TMB: 76%; OIU: 51%), $F(1, 44) = 5.84$, $p < .0$, hut not at 3;7 (TMB: 65%; OIU: 57%). The groups performed reliably above chance, minimum $t(11) = 3.96$, $p < .005$, except for the 4;2-year-old OIU group, $t(11) = 2.17$, $p = .05$. Finally, twice as many children met criterion in the TMB group (54%) as in the OIU group (25%), $\chi^2(n = .48) = 4.27$, $p < .05$.

There was a main effect of location, $F(2, 88) = 4.04$, $p < .05$, $MSE = .12$, moderated by an interaction with word set, $F(2, 88) = 3.24$, $p < .05$. As in the search trials, children hearing *on*, *in*, and *under* were most correct at the top location (69%, 42%, and 52%, respectively), and children hearing *top*, *middle*, and *bottom* were most successful for the bottom location (67, 60, and 85%, respectively). In contrast to their pattern in the search task, children did not show a tendency towards reverse object errors (49% reverse object errors).

5.3. Discussion

As predicted, children performed better on the cross-mapped box task when they heard *top*, *middle*, *bottom* than when they heard *on*, *in*, *under*. The difference is quite striking: both age groups performed well above chance when given *top*, *middle*, *bottom*, but performed at chance when given *on*, *in*, *under*. The chance performance of 3;8 and 4;2-year-olds given *on*, *in*, *under* here is consistent with the finding of chance performance for 4;1-year-olds given *on*, *in*, *under* in Experiment 3. The advantage of *top*, *middle*, *bottom* also extended to the retrieval trials: children receiving these terms were better able to remember the original hiding location than those who received *on*, *in*, *under*. In sum, we suggest that *top*, *middle*, *bottom* invited a representation that was robust enough to allow young children to preserve the relational structure despite competing object matches.

5.3.1. Alternative accounts

The non-semantic accounts of the language effect discussed earlier—surface cuing, task engagement, and verbal bypass—do not provide a basis for differentiating between the two word sets. Indeed, one might have supposed that *on*, *in*, and *under* would have a cuing advantage, as they are learned earlier and are more frequent than are *top*, *middle*, and *bottom*. However, one factor that could favor *top*, *middle*, *bottom* is differential ambiguity: the *on*, *in*, *under* word set may be more ambiguous here than the *top*, *middle*, *bottom* set. For example, children might have equated "on the box" with "on the (inside) shelf." Indeed, in the comprehension test discussed in the Introduction, 3;2-year-old children comprehended *on* less accurately than the other five spatial terms (60% vs. 80–90% correct). However, by 3;8 years of age, *on* was comprehended at 84% correct. Further, if the problem were that *on* is ambiguous, then children hearing *on*, *in*, and *under* should have performed especially poorly when given *on* (because they should often have erred towards the middle). Instead, children performed *best* at the *on* location (the top location).

Overall, we cannot rule out differential ambiguity as a contributing factor in the superiority of *top*, *middle*, *bottom* over *on*, *in*, *under*. However, evidence from an unpublished study by Loewenstein and Gentner supports our claim that relational depth is important here. Using a similar task, we contrasted two sets of location names: one with a presumed monotonic ordering (*attic*, *family room*, *basement*) and one with no such vertical ordering (*den*, *family room*, *kitchen*). Four-year-olds performed better on the monotonic set, analogous to children's superior performance with *top middle bottom*. Taken together, the results suggest that connected relational structure contributes to children's analogical performance.

6. Modeling the language effect

Our theoretical account assumes: (1) that hearing spatial relational language prompts children to encode the relations implied by the semantics of the set of terms; (2) that *top*, *middle*, *bottom* conveys a deep connected relational system, whereas *on*, *in*, and *under* each convey separate relations between the located figure and the ground; and (3) that deep relational structures—those governed by higher-order relations—are more likely to prevail in a competition against object matches than are shallow representations. We flesh out this account, based on structure-mapping theory, by describing simulation studies that embody these theoretical assumptions.

Another important motivation for carrying out the simulation is to clarify the developmental theory. Developmental gains—such as becoming able to comprehend more complex analogies—are often attributed to a change in children's basic competence, such as an increase in their mode of processing or their processing capacity. We are suggesting, on the contrary, that the rather dramatic differences in analogical mapping performance across our studies are due to changes in representation invited by the spatial relational terms. The simulations serve as a test of whether this account is sufficient to produce the pattern of data we found.

These simulations used the structure-mapping engine (SME) (Falkenhainer, Forbus, & Gentner, 1989), a computational model of the structure-mapping process that has been used to model comparison processing across a variety of tasks (e.g., Gentner, Rattermann, Markman, & Kotovsky, 1995; Markman & Gentner, 1993). We use SME without making any adjustments to its standard processing parameters. Briefly, SME takes as input two structured representations (*base* and *target*) and uses a local-to-global matching algorithm to produce one or a few overall mappings. Each mapping consists of a set of *correspondences* between the elements and predicates in the base and those in the target, and typically some *candidate inferences*—surmises about the target made by projecting information that is connected to the common system in the base. The correspondences are constrained by structural consistency and systematicity. *Structural consistency* includes: (a) the *1:1 constraint*—that each item in the base corresponds to at most one item in the target and vice versa; and (b) the *parallel co n nectivity constraint*—that if two predicates are placed in correspondence, then their arguments must also correspond. The *systematicity* principle states that deep systems of relations constrained by higher-order constraining relations are favored in the mapping over shallow interpretations containing equal numbers of relations. SME scores mappings based on the size and systematicity of their structural match. We assume that this basic cognitive mapping process is operating in all cases—both for analogy and for ordinary similarity. (See Forbus & Oblinger, 1990; Forbus, Gentner, & Law, 1995; for details.)

To avoid arbitrariness, we examined a large range of representations—six levels of relational understanding crossed with seven levels of featural richness for each of the two mapping tasks used in Experiments 1–4. As the results showed a consistent pattern across these choices, we describe a representative subset of the data. To represent the objects, we examine two levels of featural richness: for the neutral objects task, two attributes, uniform across the objects within each box; and for the cross-mapped task, a rich set of six attributes, distinctive to each pair of objects. To represent the hypothesized effects of language, we examine three levels of spatial relational representation: no relations beyond general propinquity; first-order relations; and higher-order relational structure (see the Appendix). The logic is to compare the model's predictions given the hypothesized representations with children's performance.

Table 4 shows the mapping scores. To avoid making unwarranted scaling assumptions, we use only the model's ordinal results across conditions in this comparison. On both the neutral and cross-mapped task, when the representations lacked spatial relational structure, the relational match score was no higher than the non-relational score. On the neutral objects task, once first-order relations were represented the relational choice won out. On the cross-mapped task, it was only when higher-order relations were represented that the relational match won out over the object match.

How do these results fit children's performance? The simulations predict a word set difference on the cross-mapped task but not on the neutral objects task, as was found across our studies. The simulation studies suggest that the neutral objects task is not sufficiently challenging to differentiate between a representation based only on first-order relations and one that also includes higher-order relational structure. Because first-order relations are sufficient for the task, children should do as well with *on*, *in*, *under* as with *top*, *middle*, *bottom.* This is consistent with our finding that $3\frac{1}{2}$-year-olds given *either* set of spatial terms performed above chance in

Table 4
Raw (base-normalized) SME scores by Task and Relation Level

	Relation Level 0 (no relations)	Relation Level 1 (first-order relations)	Relation Level 2 (higher-order relations)
Neutral objects task			
Relational choice	**0.061 (.735)**	**0.061 (.735)**	**0.197 (.900)**
Other choice	**0.061 (.735)**	0.044 (.530)	0.044 (.200)
Cross-mapped task			
Relational choice	0.048 (.347)	0.048 (.347)	**0.184 (.673)**
Object choice	**0.129 (.938)**	**0.103 (.752)**	0.112 (.408)

Note. Scores in parentheses are those normalized over the base representation to control for the size of the base (SME, like other analogical models, tends to give higher scores for larger matches).
Bold type indicates the model's top choice in a given simulation.
"Other" denotes the top rated non-relational choice; for cross-mapped simulations, the "other" choice was always the object match.

Experiment 1, while baseline children did not. Of course, this could simply reflect a ceiling effect; but the pattern is consistent with theory. By 4 years of age, children in both the baseline and the language condition succeeded in the neutral task, suggesting that by this age children encode simple figure–ground relations spontaneously. In contrast, on the cross-mapped task, the simulations suggest that a higher-order relational representation is required to make the relational mapping. Consistent with this suggestion, neither baseline children nor children hearing *on*, *in*, and *under* (even as old as 4;1) succeeded at the cross-mapping task. However, even 3;7-year-olds who heard the monotonic series *top*, *middle*, and *bottom* succeeded.

7. Experiment 5

We are suggesting that hearing spatial language induces a spatial relational encoding that helps children align the two boxes. The evidence so far supports this claim: children carry out the mapping more accurately when they have heard spatial relational terms (even if the language occurs before the actual task); and the semantics of the terms is crucial to the outcome. Now we carry out two further tests of the power of language to instill a conceptual representation: we ask (1) can children transfer the structure to a new pair of standards; and (2) can children retain the structure over a delay.

We first replicated Experiment 2, contrasting a spatial language group with a baseline group on the standard mapping task. Then we gave both groups a transfer test using new objects. Finally, and of most importance, we re-tested children on the standard mapping task 2 days after the initial session, using the same baseline instructions for both groups. In this design, the language group heard the terms only during the introductory task on the first day. If we find a language advantage on the transfer and retention tasks, this will be strong support for our claim that spatial relational language influences a conceptual representation.

7.1. Method

7.1.1. Participants

A total of 40 participants from the same population as in Experiment 1 were included in the current study. The younger children were 3;7 years old (range: 3;4–3;11) and the older children were 4;3 years old (range: 4;0–4;6). Half the children heard spatial language in an initial training session, and half were in a baseline condition, making 10 children in each Age by Condition group. Half of each group of 10 were male and half were female.

7.1.2. Design

The design was 2 (Age: 3;7 and 4;3) × 2 (Condition: Language and Baseline) × 3 (Location). Age and word set varied between-subjects and location varied within-subjects.

7.1.3. Materials and procedure

7.1.3.1. Training and initial mapping task. The materials, training and initial mapping task were as in Experiment 2. There were two boxes, one with three identical toy pigs and the other with identical toy chickens. During training all children were asked to place the toys at specified locations. Language children were trained on only one box, and were asked to place objects in locations designated by *top*, *middle*, *bottom* (e.g., "Can you put this at the top of the box?"). Baseline children saw both boxes, one at a time, and were asked to place a toy "here," guided by a finger point. All children placed toys at each location three times. After one of the two kinds of training, both groups were given the same standard neutral objects mapping task. As in Experiment 2, the "winner" had a prize inside it. No relational language was used in this or any subsequent task: all children received the same "baseline" instructions from this point onwards.

7.1.3.2. Transfer and delayed retention tasks. After completing the initial mapping task, the children were given the transfer task. They were told that they were going to play the same hiding and finding game with new objects (see Fig. 1). The new reference objects were made from plastic buckets turned upside down, with portions cut away, and a middle shelf added. They had a three-tiered spatial structure like the original boxes, but differed in color (red or yellow with large spots, instead of plain blue or white), shape (curved instead of straight), material (plastic instead of wood), and size (a few inches shorter). New toys were also used (sheep and cows instead of the initial pigs and chickens). As in the initial mapping task, during the task, the experimenter made no reference to the spatial relational terms: e.g., "I'm making this one the winner, and putting it right here." Children were shown all the items, there was a demonstration trial with an item placed off to the side, followed by six pairs of search and retrieval trials. Two days later, all children were tested again on the initial box mapping task in the same manner as in the initial session. Critically, the experimenter did not use the spatial language terms in either the transfer task or in the retention session.

7.2. Results

7.2.1. Search trials

7.2.1.1. Initial mapping task. As in the prior studies, the results showed a clear effect of spatial language on performance. Language children performed better on the initial mapping task than did children in the baseline condition, 70% versus 53% correct, $F(1,36) = 5.77$, $p < .05$, $MSE = .14$ (see Fig. 6 and Table 2). There was no effect of nor interaction with location. Older children did not perform significantly better than younger children, 68% versus 55% correct, $F(1,36) = 3.69$, $p = .06$. Planned contrasts for each age group showed a significant advantage for the language group (80%) over the baseline group (57%) among the older children, $F(1,36) = 5.65$, $p < .05$, but not among the younger children (60% vs. 50%). All groups, including the younger baseline children, performed reliably above chance (.33), $t(9) = 2.372$, $p < .05$. An analysis of the number of children who met the 5 out of 6 criterion for above-chance performance confirmed that more language (55%) than baseline (15%) children did so, $\chi^2(1, n = 40) = 7.03$, $p < .01$.

7.2.1.2. Transfer task. The younger group showed a language effect, but the older group did not. There was no main effect of age nor of location, and the effect of language was not significant (69% vs. 56% for language and baseline, respectively), $F(1,36) = 3.26$, $p = .08$, $MSE = .16$. Planned contrasts showed that the younger language group performed better than their baseline group (70% vs. 45%), $F(1,36) = 5.74$, $p < .05$; but the older language (68%) and baseline (67%) groups

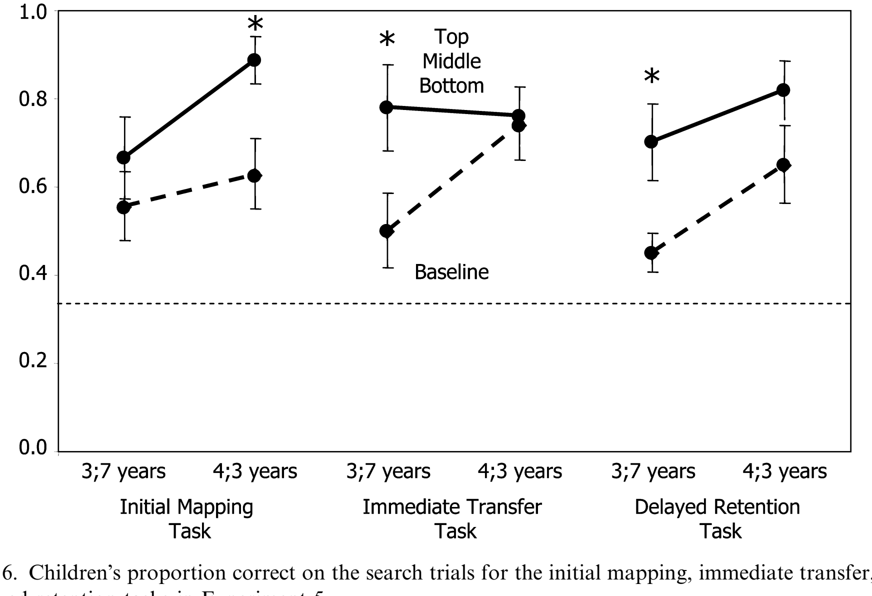

Fig. 6. Children's proportion correct on the search trials for the initial mapping, immediate transfer, and delayed retention tasks in Experiment 5.

did not differ. All groups except the younger baseline group performed above chance (33%).

7.2.1.3. Delayed retention task. The effects of language exposure were clearly evident 2 days later, even though the children were not reminded of the terms. Language children performed better than baseline children (76% vs. 55%), $F(1,36) = 8.09$, $p < .01$, $MSE = .16$. There were no effects of location. Older children performed better than younger children (73% vs. 58%), $F(1,36) = 4.68$, $p < .05$. Planned contrasts showed a significant advantage for the language group (70%) over the baseline group (45%) among the younger children, $F(1,36) = 5.83$, $p < .05$, but not the older children (82% versus 65%), $F(1,36) = 2.59$, $p = .12$. All age groups, including the younger baseline group, performed above chance (.33), $t(9) = 2.69$, $p < .05$. More language (60%) than baseline (25%) children met the criterion for above-chance performance, $\chi^2(1, n = 40) = 5.01$, $p < .05$.

7.2.2. Retrieval trials

Children's performance on the retrieval trials for the three tasks largely mirrored their performance on the search trials (see Table 2). As expected, there was an overall advantage for hearing spatial language, $F(1,36) = 10.01$, $p < .01$, $MSE = .14$; and older children (82%) performed better than younger children (57%), $F(1,36) = 12.71$, $p < .01$. Confirming this pattern, a 3 (Task: initial mapping, immediate transfer and delayed retention) $\times$ 2 (Trial Type: search and retrieval) $\times$ 2 (Age) $\times$ 2 (language and baseline) ANOVA showed (in addition to effects of age and language condition) the unsurprising effect that the retrieval trials were easier than the search trials (71% vs. 63% correct), $F(1,36) = 15.27$, $p < .001$, $MSE = .02$, and an interaction between trial type and age showing a larger age difference on the retrieval trials than the search trials, $F(1,36) = 4.85$, $p < .05$. There were no effects of task nor any interactions involving task. All children performed above chance on the retrieval trials, minimum $t(9) = 2.40$, $p < .05$.

7.3. Discussion

In this study we tested a key aspect of the claim that language can influence cognition—whether the effects of hearing a set of semantic descriptors can persist over time. We found that exposure to spatial relational terms provided an immediate advantage on the mapping task, as in the prior studies. Critically, this mapping advantage extended over time: children who heard the spatial terms still performed better than their baseline counterparts 2 days later. The 3;7-year-olds also showed a language advantage on a transfer task. Spatial language was used only during the placement task at the outset—not during the three mapping tasks. Thus hearing spatial language provided more than a localized and temporary mapping strategy. It appears to have induced a shift in how scenes were represented that was sufficiently durable to provide a sustained advantage on challenging spatial mapping tasks.

8. General discussion

We tested the hypothesis that hearing language for spatial relations would help children to encode relational structure and to carry out spatial analogies—a specific instance of our general hypothesis that relational language fosters developing representational tools for engaging in relational thinking (Gentner, 2003; Gentner & Loewenstein, 2002). In five experiments, we consistently found that preschool children performed better on spatial relational mapping tasks if they had heard spatial relational terms applied to the task situation than if they did not, confirming our first prediction. Our second prediction was also borne out: the advantage of spatial relational language appeared at later ages for more difficult tasks. Children hearing *on*, *in*, and *under* showed a language advantage on the neutral objects task by 3;7 years (Experiments 1 and 2). On the more challenging cross-mapped task, this language advantage did not appear until 4:7 (Experiment 3). The third prediction was that the effects of spatial relational language would be specific to the semantics of the terms used. This was borne out in Experiment 4, in which we contrasted two sets of spatial relational terms on the cross-mapped task, one of which (*top–middle–bottom*) conveys a more deeply connected system of relations than the other (*on–in–under*). The results were dramatic: whereas children who heard *on–in–under* achieved above-chance performance only at 4;7, children who heard *top–middle–bottom* performed above chance at 3;7—a full year earlier. Simulations confirmed that the structure-mapping process, combined with the representational account we propose, can generate the complex patterns of interaction between language support, task, and age found here.

In the final study, we asked whether relational language could induce an enduring change in representation. We found that the effects of spatial relational language persisted over time: children tested 2 days after initial exposure to spatial language terms still showed an advantage relative to children who did not hear the spatial terms. Further, at least among younger children, those who had heard a relational description were better able to transfer the analogy to new materials. Taken together, the evidence suggests that hearing spatial language leads to richer encoding of spatial relations and thus to better analogical mapping.

Alterative accounts. We considered three alternative accounts of the language effect—task engagement, surface cuing, and verbal bypass—and found that none of them can plausibly account for the overall results. None of these accounts predicts the observed advantage of *top middle bottom* over *on in under* in Experiment 4, nor the nature of children's comments about the task in Experiment 1. They also cannot explain children's success in Experiments 2 and 5, in which language was only applied to one of the boxes, and was not used during the actual mapping task. Finally, none of the alterative strategies can predict our finding that the effects of one brief language session were still apparent after a 2-day delay (in Experiment 5). Thus we conclude that hearing the spatial language induced a conceptual representation of spatial relations.

However, we are not suggesting that the effect of language is only to invite a particular relational representation. It seems likely that language operates in several ways. Language appears to support cognitive flexibility (Jacques & Zelazo, in press) and to guide children in their interpretation of the task (Gelman & Greeno, 1989). For example, spatial language might have drawn children's attention generally towards spatial information, as well as inducing the specific encodings invited by the terms.

8.1. Related work

Our results are consistent with the pattern found by Gentner and Rattermann (1991; Rattermann & Gentner, 1998b). They found effects of introducing relational language on 3-year-old children's ability to carry out a relational mapping. In these studies, the relational pattern was monotonic increase in size across a line of objects; the correct answer was based on matching relative size and position. As in the present studies, the mapping was made difficult by introducing a cross-mapping between the object matches and the relational correspondences. The results showed that children who heard language conveying a monotonic relational structure (either *big–little–tiny* or *Daddy–Mommy–Baby*) performed far better than those who did not. These findings are further evidence for a facilitating effect of relational language on children's appreciation of relational similarities.

The language advantage is not restricted to the period of initial acquisition. We found benefits among 5-year-olds—at least 18 months after children comprehend the terms. Indeed, there is evidence that even adults may make use of internal spatial language if the task is sufficiently challenging. For example, Wolff, Vassilieva, and Burgos (2002) gave people a mental rotation task involving spatial scenes. For many configurations, people showed the typical pattern found in mental rotation studies: the reaction times increased with the degree of rotation (Shepard & Cooper, 1982). However, when the spatial relations could readily be labeled, people were able to "shortcut" the task: they showed a fast, flat reaction time pattern, suggesting that they solved the task by matching the relational descriptions instead of by mental rotation. A complementary result is the decrement in spatial reasoning found when adults were prevented from using spatial language, as discussed earlier (Hermer-Vasquez et al., 1999).

8.2. Extensions and implications

Our research fits within a larger tradition of work on developmental interactions between language and cognition (Nelson, 1996). For example, young children's willingness to make inductive inferences between entities is enhanced by the presence of a common label (Gelman & Markman, 1987; Waxman, Lynch, Casey, & Baer, 1997). Further, differences in language-specific semantic patterns can lead to differences in children's very early patterns of semantic extension (Bowerman & Choi, 2001; Imai & Gentner, 1997). However, with a few exceptions (Gopnik & Choi, 1990; Smith & Sera, 1992), most previous work on the effects of labels on cognitive development has focused on the effects of noun labels on concepts of objects and entities. Our results suggest that the effects of language on conceptual development also hold for relational language. Indeed, following Gentner's relational relativity conjecture, according to which relational terms vary cross-linguistically to a greater degree than noun terms (Gentner, 1982; Gentner & Kurtz, in press; Gentner & Boroditsky, 2001), we conjecture that effects of habitual language on thought are more likely with relational terms than with concrete nouns.

8.3. Conclusion

The current studies shed light on a potentially powerful force in learning to notice and use relational structure. Hearing words that name spatial relations facilitated children's encoding and mapping of spatial relations. If indeed relational language generally invites noticing and using relations, then the acquisition of relational language is instrumental in the development of abstract thought.

Appendix

We gave SME representations like those hypothesized for the children as a test of whether these representational changes, together with the processing assumptions of structure-mapping theory, are sufficient to generate the obtained pattern of results across the experimental conditions. Our goal here was not to model language learning per se, but rather to model the effects of using different linguistically conveyed spatial representations. SME's representations include three kinds of elements relevant here. First are *entities*—the objects or characters in a domain (e.g., box or card). Second are *attributes*—unary predicates that describe object features or properties: e.g., BLUE[box] or FLAT[card]. Third are *relations*—multi-place predicates that link entities, attributes or other relations: e.g., ON[card, box] or ABOVE[card 1, card 2]. *First-order relations* take entities or attributes as arguments (as just shown), whereas *higher-order relations* take other relations as arguments: for example, MONOTONIC [ABOVE(card1, card2), ABOVE(card2, card3)], where *MONOTONIC* [*R* (*1*, *2*) *R* (*2*, *3*)] is a higher-order relation conveying an ordering relation such that R (1, 2) and R(2, 3) imply R (1, 3).

Two orthogonal factors must enter into the simulation: (1) the two tasks — neutral vs. cross-mapped—for require representing different levels of featural richness; and (2) the presumed effects of hearing spatial terms, which require representing different kinds of relational representation. To avoid arbitrary assumptions, we ran six levels of relational representation and seven levels of featural richness. As all of these showed the same basic pattern, for brevity we show two levels of featural richness and three levels of relational representation (see the tables). For the neutral objects task, the objects were simple colored cards, identical within each box and differing in one attribute across boxes. For example, the attributes GRAY and RECTANGLE were used for all three cards in the hiding box, and BLUE and RECTANGLE were used for all three cards in the finding box. For the cross-mapped task, three pairs of cards sharing a rich, distinctive set of features were represented in different relative positions at the two boxes. To simulate children's relational understanding, we generated three levels of representations corresponding to the three levels of language support: no language, *on–in–under* (OIU), and *top–middle–bottom* (TMB). The no-language level had three identical general locative relations—NEAR(a, b). The OIU level had three distinct first-order relations—ON(a, b), IN(a, c), UNDER(a, d). The TMB level had the three first-order relations as well as a higher-order relation: MONOTONIC {ABOVE(a, b), ABOVE (b, c)}. SME was given the six representational pairs that result from crossing the three relational levels with the two feature sets.

SME operates entirely in literal similarity mode; there is no special mode for finding analogies. When there are competing bases for mapping (such as object matches vs. relational matches) SME computes more than one interpretation and provides a match score for each interpretation based on the size and depth of the common structure. No experimenter intervention occurs during its runs, nor is the order of operations specified by the experimenter. All simulations here used the standard mapping parameters that have been used in prior studies of adults and children (see Forbus et al., 1995).

For both the neutral-objects and the cross-mapped task, the question is under what conditions SME will form the relational mapping and prefer it over alternative mappings—e.g., the object-based mapping in the case of the cross-mapped task. The question was whether its pattern of selecting the relational mapping would be consistent with that of the children. The results, shown in Table 4, bear this out.

Representations used in the simulations (in LISP notation), showing (a) the three levels of relational representation; and (b) the objects used in the neutral and cross-mapped tasks.

(a) Representation of the relations

Relation Level (representing)	Representation
1 (no language)	(near box card-1) (near box card-2) (near box card-3) (behind card-1 star)
2 (on, in, under)	(on box card-1) (in box card-2) (under box card-3) (behind card-1 star)
3 (top, middle, bottom)	(top box card-1) (middle box card-2) (bottom box card-3) (behind card-1 star) (monotonic-height (above card-1 card-2) (above card-2 card-3))

(b) Representations of the objects in the neutral and cross-mapped tasks

Task	Card	Hiding box	Finding box
Neutral objects task	1	(gray card-1)	(blue card-1)
		(rectangle card-1)	(rectangle card-1)
	2	(gray card-2)	(blue card-2)
		(rectangle card-2)	(rectangle card-2)
	3	(gray card-3)	(blue card-3)
		(rectangle card-4)	(rectangle card-4)
Cross-mapped task	1	(black-back card-1)	(green-back card-1)
		(dotted-back card-1)	(smooth-back card-1)
		(circle-obj card-1)	(rectangle-obj card-1)
		(small-obj card-1)	(big-obj card-1)
		(purple-obj card-1)	(yellow-obj card-1)
		(plant-cat card-1)	(animal-cat card-1)
	2	(green-back card-2)	(red-back card-2)
		(smooth-back card-2)	(checkered-back card-2)
		(rectangle-obj card-2)	(triangle-obj card-2)
		(big-obj card-2)	(medium-obj card-2)
		(yellow-obj card-2)	(brown-obj card-2)
		(animal-cat card-2)	(food-cat card-2)
	3	(red-back card-3)	(black-back card-3)
		(checkered-back card-3)	(dotted-back card-3)
		(triangle-obj card-3)	(circle-obj card-3)
		(medium-obj card-3)	(small-obj card-3)
		(brown-obj card-3)	(purple-obj card-3)
		(food-cat card-3)	(plant-cat card-3)

References

American Heritage Dictionary of the English Language (2000). New York: Bartleby.com.

Blades, M., & Cooke, Z. (1994). Young children's ability to understand a model as a spatial representation. *Journal of Genetic Psychology, 155*(2), 201–218.

Bluestein, M., & Acredolo, L. (1979). Developmental changes in map reading skill. *Child Development, 50*, 691–697.

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press.

Bowerman, M., & Levinson, S. (2001). *Language acquisition and conceptual development*. Cambridge, UK: Cambridge University Press.

Bowerman, M., & Pederson, E., (1992, December). *Crosslinguistic perspectives on topological spatial relationships*. Paper presented at the annual meeting of the American Anthropological Association, San Francisco, CA.

Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369–412). New York: Cambridge University Press.

Chen, Z., Sanchez, R. P., & Campbell, T. (1997). From beyond to within their grasp: The rudiments of analogical problem solving in 10- and 13-month-olds. *Developmental Psychology, 33*(5), 790–801.

Clark, E. V. (1980). Here's the top: Nonlinguistic strategies in the acquisition of orientational terms. *Child Development, 51*(2), 329–338.

Clement, C., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science, 15*, 89–132.

Coventry, K. R. (2001). *Spatial language: Cognitive and computational perspectives*. Dordrecht: Kluwer.

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science, 238*, 1556–1557.

DeLoache, J. S. (1995). Early symbol understanding and use. In D. Medin (Ed.). *The psychology of learning and motivation* (Vol. 32). New York: Academic Press.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*, 1–63.

Feist, M. I. (2000). On in and on: An investigation into the linguistic encoding of spatial scenes. *Dissertation Abstracts International, 61*(6), 2280A.

Forbus, K. D., & Oblinger, D. (1990). Making SME greedy and pragmatic. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, pp. 61–68.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*(2), 141–205.

Gattis, M. (Ed.). (2001). *Spatial schemas and abstract thought*. Cambridge, MA: MIT Press.

Gelman, R., & Greeno, J. G. (1989). On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.), *Knowing and learning: Essays in honor of Robert Glaser* (pp. 125–186). Hillsdale, NJ: Erlbaum Associates.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532–1541.

Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory, 4*(2), 161–178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D. (1988). Metaphor as structure-mapping: The relational shift. *Child Development, 59*, 47–59.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition*. Cambridge, MA: MIT Press.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge, UK: Cambridge University Press.

Gentner, D., & Goldin-Meadow, S. (Eds.). (2003a). *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.

Gentner, D., & Goldin-Meadow, S. (2003b). Whither Whorf. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 3–14). Cambridge, MA: MIT Press.

Gentner, D., & Kurtz, K. J. (in press). Relational categories. In W. K. Ahn, R. L. Goldstone, A. B. Markman, B. C. Love, & P. W. Wolff (Eds.), *Categorization inside and outside the lab: Essays in honor of Douglas L. Medin*. Washington, DC: APA.

Gentner, D., & Loewenstein, J. (2002). Relational thinking and relational language. In J. Byrnes & E. Amsel (Eds.), *Language, literacy, and cognitive development*. Hillsdale, NJ: Erlbaum.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Brynes (Eds.), *Perspective on thought and language: Interrelations in development* (pp. 225–277). New York: Cambridge University Press.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300.

Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: Evidence for two systems in processing space-time metaphors. *Language and Cognitive Processes, 17*, 537–565.

Gentner, D., Rattermann, M. J., Markman, A. B., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263–313). Hillsdale, NJ: Erlbaum.

Gobbo, C., & Chi, M. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development, 1*(3), 221–237.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23–45.

Gopnik, A., & Choi, S. (1990). Do linguistic differences lead to cognitive differences? A cross-linguistic study of semantic and cognitive development. *First Language, 10*, 199–215.

Gumperz, J. J., & Levinson, S. C. (1996). *Rethinking linguistic relativity*. Cambridge, MA: Cambridge University Press.

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.

Hall, D. G., & Waxman, S. R. (1993). Assumptions about word meaning: Individuation and basic-level kinds. *Child Development, 64*(5), 1224–1241.

Hermer-Vasquez, 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.

Hermer-Vasquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology, 39*, 3–36.

Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. Cambridge, UK: Cambridge University Press.

Holland, J. H., Holyoak, K. F., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review, 98*, 377–389.

Imai, M., & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62*, 169–200.

Jacques, S., & Zelazo, P.D. (in press). Language and the development of cognitive flexibility: Implications for theory of mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind.* Oxford: Oxford University Press.

Johnston, J. R. (1988). Children's verbal representation of spatial location. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (Eds.), *Spatial cognition: Brain bases and development* (pp. 195–205). Hillsdale, NJ: Lawrence Erlbaum.

Kuczaj, S. A., & Maratsos, M. (1975). On the acquisition of front, back and side. *Child Development, 46*, 202–210.

Kuenne, M. R. (1946). Experimental investigation of the relation of language to transposition behavior in young children. *Journal of Experimental Psychology, 36*(6), 471–490.

Levinson, S. C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: Language affects spatial reasoning. *Cognition, 84*(2), 155–188.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83*, 265–294.

Liben, L. S. (1998). Developing an understanding of external spatial representations. In I. E. Sigel (Ed.), *The development of representational thought: Theoretical perspectives*. Hillsdale, NJ: Erlbaum.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition & Development, 2*(2), 189–219.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom & M. A. Peterson (Eds.), *Language and space: Language speech, and communication*. Cambridge, MA: MIT Press.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*, 431–467.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16*(1), 1–27.

Meints, K., Plunkett, K., Harris, P. L., & Dimmock, D. (2002). What is 'on' and 'under' for 15-, 18- and 24-month-olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology, 20*(1), 113–130.

Nelson, K. (1996). *Language in cognitive development: Emergence of the mediated mind*. New York, NY, USA: Cambridge University Press.

Newcombe, N. S. (2002). The nativist–empiricist controversy in the context of recent research on spatial and quantitative development. *Psychological Science, 13*(5), 395–401.

Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning*. Cambridge, MA, USA: The MIT Press.

Presson, C. C. (1982). The development of map-reading skills. *Child Development, 53*(1), 196–199.

Ramscar, M., & Pain, H. (1996). Can a real distinction be made between cognitive theories of analogy and categorisation? *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*. (pp. 346–351). New Jersey: Lawrence Earlbaum Associates.

Rattermann, M. J., & Gentner, D. (1998a). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development, 13*, 453–478.

Rattermann, M. J., & Gentner, D. (1998b). The effect of language on similarity: The use of relational labels improves young children's performance in a mapping task. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory & data from the cognitive, computational, and neural sciences* (pp. 274–282). Sophia: New Bulgarian University.

Regier, T., & Carlson-Radvansky, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology, 130*(2), 273–298.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology, 13*(4), 629–639.

Shepard, R., & Cooper, L. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.

Smith, L. B. (1984). Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. *Child Development, 55*, 363–380.

Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology, 24*(1), 99–142.

Talmy, L. (1983). How language structures space. In H. Pick & L. Acredolo (Eds.), *Spatial orientation: Theory, research, and application*. New York: Plenum.

Tomasello, M. (1987). Learning to use prepositions: A case study. *Journal of Child Language, 14*(1), 79–98.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press, Original work published (1934).

Waxman, S. R., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1*, 139–156.

Waxman, S. R., Lynch, E. B., Casey, K. L., & Baer, L. (1997). Setters and samoyeds: The emergence of subordinate level categories as a basis for inductive inference in preschool-age children. *Developmental Psychology, 33*, 1074–1090.

Wilcox, S., & Palermo, D. S. (1980). "In," "on," and "under" revisited. *Cognition, 3*(3), 245–254.

Wolff, P., Vassilieva, T., Burgos, F., (May 2–4, 2002). *Language and mental rotation in the localization of objects*. 2002 Annual Meeting of the Midwestern Psychological Association, Chicago, IL.

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science, 7*, 121–126.