Cognitive Development 50 (2019) 177–194

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Finding the *middle*: Spatial language and spatial reasoning

Nina K. Simms[a,*], Dedre Gentner[b]

[a] Spatial Intelligence and Learning Center, Northwestern University, 2120 Campus Drive, Suite 162, Evanston, IL, 60208-2710, USA
[b] Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL, 60208-2710, USA

ARTICLE INFO

Keywords:
Midpoint
Spatial cognition
Spatial reasoning
Language and thought

ABSTRACT

Learning relational language has been implicated in the development of spatial relational reasoning. We hypothesized that children's encoding of the midpoint, a complex spatial relation, would be predicted by their knowledge of the relevant spatial terms, *middle* and *between*. Children aged 2½ to 5½ were asked to find an object hidden at the midpoint between two landmarks, which varied in their location and distance. Children improved with age on the search task; further, children's performance after feedback suggested that encoding the midpoint relation poses a genuine challenge for young children. Children's knowledge of the words *middle* and *between,* assessed in a separate task, predicted their search success beyond what was predicted by age or knowledge of other spatial terms. These findings are consistent with the view that relevant spatial word knowledge supports the representation and use of this complex spatial concept, a proposal that future work will seek to address.

1. Introduction

Spatial thinking is central to both human and non-human cognition, and humans have developed and use elaborate symbolic systems like language with which to communicate and reason about space. The role of language in spatial thinking has become a central question, particularly in areas where our abilities diverge sharply from other species without such symbol systems (e.g., Haun, Rapold, Call, Janzen, & Levinson, 2006). Spatial reasoning abilities also develop significantly during childhood (Newcombe & Huttenlocher, 2003; Newcombe, Uttal, & Sauter, 2013), as does children's spatial language knowledge (e.g. Johnston, 1988). In this paper, we explore the relationship between children's spatial language and their spatial reasoning for a skill that shows dramatic differences between human adults and other species, and human adults and children: encoding the midpoint. We begin by reviewing research in the development of midpoint encoding before turning to potential effects of language.

1.1. Midpoint and relational encoding

The midpoint is an interesting and potentially challenging spatial relation in that it encodes location with respect to two or more ground objects or landmarks. Many spatial relations involve a relationship between an object (the figure) and one other object (the ground). For instance, one object might be *near* another. More specifically, but still involving only a single ground object, a figure could be *inside* or *behind* another. In contrast, the midpoint relation requires understanding (a) that the figure lies *between* two ground objects, and also (b) that it is *equidistant* from both. In other words, the midpoint is defined by its proportional relationship to a configuration. As such, midpoint encoding is an example of *relational coding* – a type of spatial encoding in which location is encoded

* Corresponding author.
E-mail addresses: ninasimms@northwestern.edu (N.K. Simms), gentner@northwestern.edu (D. Gentner).

https://doi.org/10.1016/j.cogdev.2019.04.002
Received 2 November 2018; Received in revised form 3 March 2019; Accepted 21 April 2019
Available online 14 May 2019

with respect to the relations among multiple landmarks, rather than a single landmark (e.g., MacDonald, Spetch, Kelly, & Cheng, 2004; Uttal, Sandstrom, & Newcombe, 2006).[1]

A compelling illustration of the challenge of encoding the midpoint relation comes from a study with gerbils (Collett, Cartwright, & Smith, 1986). The animals were trained to find seeds in the middle of two landmarks, always the same distance apart. When the landmarks were expanded to a wider distance, instead of searching in the middle of them – which would have preserved the proportional relationships to each landmark – the gerbils searched at a location that preserved the training distance from only one of the landmarks. That is, the gerbils had not encoded the location as *in the middle of two landmarks*, but rather had encoded it as a specific distance from a single landmark.

In fact, successful encoding of the midpoint relation has been difficult or impossible to demonstrate in many other animals, including fish (Sovrano, Bisazza, & Vallortigara, 2007), a variety of birds (Kamil & Jones, 2000; Kamil & Jones, 1997; Spetch, Cheng, & MacDonald, 1996; Spetch et al., 1997), and other primates (Potì et al., 2010; Sutton, Olthof, & Roberts, 2000). For animals that have demonstrated midpoint-encoding strategies, extensive training has generally been required (e.g., Jones, Antoniadis, Shettleworth, & Kamil, 2002; Potì et al., 2010). In contrast, human adults fluently and preferentially use relational coding for midpoint arrays (MacDonald et al., 2004; Spetch et al., 1996, 1997). For example, MacDonald and colleagues trained undergraduates to find a token hidden at the midpoint of four landmarks. At test, the four landmarks were expanded to a larger distance, and participants were again asked to find the hidden token. All participants searched for the token at the center of the larger array, suggesting that they had encoded the location with respect to all four of the landmarks and the relationships among them. Marmoset monkeys given the same task, however, concentrated their searches near one of the four landmarks, indicating that they (like the gerbils described above) had encoded location with respect to only a single landmark.

1.2. Development of relational encoding

As might be expected by its relative complexity, midpoint encoding emerges later than single-landmark encoding strategies in development. By one-year-old, children can successfully locate objects directly at a single landmark (Bushnell, McKenzie, Lawrence, & Connell, 1995). This type of strategy is known as a *beacon* strategy and roughly corresponds to encoding location as *at*, *by*, or *near* a landmark. By 2-years-old, children can locate an object indirectly with respect to a single landmark by encoding the specific distance and direction from that landmark (known as *vector* coding) (Bushnell et al., 1995; DeLoache & Brown, 1983; Newcombe, Huttenlocher, Drummey, & Wiley, 1998), although children may not always recognize the relevance of landmarks for encoding location in this way (Shusterman, Lee, & Spelke, 2011). However, successful relational coding, which has been most often explored using midpoint configurations, has not been demonstrated until at least 3-years-old (Ankowski, Thom, Sandhofer, & Blaisdell, 2012). In one ingenious exploration, Uttal et al. (2006) found that 4-to-5-year-olds could pass a close analog of the task used with gerbils described earlier (Collett et al., 1986). They tested the children in a large field, preserving the same relative scale between the children and the array as the gerbils had experienced. In contrast to the gerbils, the children searched unanimously at the midpoint. However, the age at which children pass midpoint tasks has varied widely across studies; in some studies even 7- to 9-year-olds still do not use a midpoint strategy (MacDonald et al., 2004).

Studies of midpoint encoding have varied many details (e.g., scale) that may account for the differences in ages reported to succeed. Despite this variability, these investigations have all used the same basic paradigm. In order to differentiate children's encoding with respect to a single landmark versus encoding with respect to a configuration, children are first trained to criterion on arrays with a constant inter-landmark distance; then the array is expanded to a larger distance, and children's spontaneous search strategies in that new array are assessed with a single trial.[2] These investigations converge in suggesting that children become increasingly likely to encode the midpoint relation with age (Ankowski et al., 2012; Spetch & Parent, 2006).

Some evidence suggests that children's understanding of midpoint develops in a piecemeal fashion, with children encoding certain aspects of the complex relation before others. Specifically, children's error patterns on midpoint search tasks suggest that they may encode the *between* component before incorporating *equidistance*. For example, Spetch and Parent (2006) trained 3- to 5-year-old children to find a hidden toy in the middle of two landmarks in a row of boxes. During training, many of the children searched in a location adjacent to one of the landmarks (rather than in the middle). Strikingly, for the younger children, these searches occurred equally often between and outside the landmarks. In contrast, older children rarely searched outside the landmarks, suggesting that they understood the relevance of both landmarks (the *between* aspect) but were not precisely encoding location at the midpoint. This delay in precision is perhaps not surprising given that a key aspect of the midpoint relation is its proportional relationships, and children become increasingly precise in their proportional understanding with age (e.g., first being able to discriminate qualitatively different proportions from one another before being able to make finer distinctions; Spinillo & Bryant, 1991). However, the prevalence of outside errors varies widely across studies and ages (e.g., MacDonald et al., 2004; Uttal et al., 2006), so more work is needed to explore the trajectory of this pattern.

Children also exhibit flexibility in their strategies for encoding location (Marsh, Spetch, & MacDonald, 2011; Uttal et al., 2006). The 4- and 5-year-old children in Uttal et al. (2006) study who were trained on two-landmark arrays all used a midpoint strategy, but

[1] While it could be argued that any location encoded with respect to another landmark or object is relational, in that it encodes a relationship between a figure and ground, the convention in comparative research on spatial tasks reserves "relational coding" for cases involving relations among multiple lower-order relations – for example, whether an object's distance from landmark A is equivalent to its distance from landmark B.

[2] But see Uttal et al. (2006) for an exception to the single test-trial design.

they were able to switch to a vector strategy when subsequently presented with only a single landmark. Conversely, children trained on single-landmark arrays were able to successfully switch to a midpoint strategy immediately after finding an object in the middle of a two-landmark configuration. This flexibility is evident in children as young as those in Uttal and colleagues' study, but children's flexibility also increases with age. Marsh et al. (2011) found that 8- to 12-year-olds were more likely than 3- to 7-year-olds to try variable strategies when their initial strategies proved unsuccessful.

By adulthood, humans preferentially adopt a midpoint strategy to encode location, in contrast to non-human animals (MacDonald et al., 2004; Spetch et al., 1996, 1997). But what underpins this developmental trajectory?

1.3. Spatial language and spatial thinking

Our central hypothesis is that acquiring and using terms like *middle* that refer to the midpoint relation is instrumental in the development of midpoint encoding. Language may support spatial thinking in a variety of ways. For instance, linguistic descriptions can help children understand what kind of information, including landmarks, might be relevant on a spatial task (Loewenstein & Gentner, 2005; Shusterman et al., 2011). Here, we explore a more specific connection between spatial language and spatial thinking. Gentner and colleagues have proposed that acquiring relational language fosters children's understanding of the corresponding relational concepts (2010, Gentner, 2016; Gentner, 2003; Gentner & Christie, 2012; Gentner & Namy, 2006). According to this proposal, language supports relational learning in part because repeated application of the term to instances of the relation prompts children to compare the instances and thereby derive their common structure. For example, Casasola and colleagues have found that hearing labels for spatial relations – but not other kinds of language – can help young children learn spatial relational categories (Casasola, 2005; Casasola & Bhagwat, 2007; Casasola, Bhagwat, & Burke, 2009). A further way in which linguistic labels support children's thinking is that a label invites unifying and preserving the named conceptual structure; this renders it more accessible for future use (Gentner, 2003, 2010; Gentner, Özyürek, Gürcanli, & Goldin-Meadow, 2013; Lupyan, Rakison, & McClelland, 2007; Son, Smith, Goldstone, & Leslie, 2012).

We propose that learning and using the spatial term *middle* fosters encoding and reasoning about the midpoint relation via these mechanisms. As children learn the meaning of the word *middle*, they also refine their understanding of the midpoint relation by comparing multiple instances. Storing the abstracted relation along with the label *middle* could make the representation more stable, efficient, and readily accessible for encoding subsequent midpoint examples. Children may also be able to engage verbal strategies – for instance, describing something as *in the middle* – to help support midpoint encoding.

It is possible that acquiring the term *between*, which captures important components of midpoint like the reference to multiple objects, may also confer some benefit to midpoint encoding. While *between* does not convey the equidistance that is critical for the midpoint relation, children with knowledge of *between* may be able to recruit the associated relational concept to gain purchase on the more precise midpoint relation. A child who notes that a location falls between two objects may be in a better position to adjust their encoding to incorporate equidistance than a child who has not even captured both objects in their spatial encoding.

There is considerable evidence that learning spatial language supports children's reasoning about spatial relations. For example, early spatial vocabulary knowledge predicts performance on nonverbal spatial reasoning tasks, both at the same time (Balcomb, Newcombe, & Ferrara, 2011) and at later points in development (Pruden, Levine, & Huttenlocher, 2011). There is also evidence for benefits of specific spatial terms. First, hearing relevant spatial language during challenging spatial tasks has been shown to improve children's spatial reasoning performance (Loewenstein & Gentner, 2005; Miller, Patterson, & Simmering, 2016). For example, Loewenstein and Gentner found that preschoolers who heard labels for spatial locations (such as *top*, *middle*, or *bottom*) in a spatial mapping task were more likely to correctly rely on a spatial alignment, and ignore distracting featural information, than children who did not—consistent with the idea that hearing spatial language may benefit performance by inviting a particular construal of the space.

Second, and more relevant to the present hypothesis, there is evidence that knowledge of specific spatial terms predicts children's ability in nonlinguistic tasks that rely on the corresponding relations. For example, children's ability to produce the spatial terms *left* and *right* (but not other spatial terms or cognitive factors) predicted their ability to use a salient landmark to reorient themselves in space in order to locate a hidden object that is left of or right of the landmark (Hermer-Vazquez, Moffet, & Munkholm, 2001), even though the terms were not used during the task. Our research asks whether this connection exists for the midpoint relation as well.

Some prior studies have found hints of a link between midpoint search performance and children's relevant spatial word knowledge. For example, MacDonald et al. (2004) found that most of their 3- and 4-year-olds searched near one of the landmarks in the expanded test array, suggesting they had encoded location using a single-landmark strategy. However, a few children did search in the middle of the test array – and these were the same ones who spontaneously said *middle* during the task. Spetch and Parent (2006) also reported that the few children who described the location of a hidden object using the words *middle* or *between* all used a middle rule on their search task.

Noting these suggestive patterns, Ankowski et al. (2012) carried out a pioneering exploration of the relationship between access to the word *middle* and children's midpoint search performance, using a task modeled on MacDonald et al. (2004). They found two results that suggested that access to the word *middle* supported children's midpoint encoding. First, they found that 2- to 4-year-olds who heard the experimenter describe the trained hiding location as "in the middle" were marginally more likely to search at the midpoint in the expanded array than children who heard it described as "here." Second, they assessed children's comprehension of the word *middle* using a separate task and found that children who searched at the midpoint scored higher on the *middle* assessment than children who used an incorrect strategy.

These findings are consistent with the idea that knowledge of *middle* supports midpoint search performance. However, they also raise a number of questions. First, age was related to performance on both tasks; could it be that older children simply performed better on both tasks? In addition, the relationship between comprehension and search performance was collapsed across children who heard "in the middle" on the search task and children who did not. Hearing the word during the search task could have improved comprehension performance on the subsequent language assessment—leaving open the possibility that the relationship between language and performance might have been driven by children who heard the word *middle* on both tasks.

1.4. Present study

Our goal is to test whether knowledge of particular spatial relational terms relate to children's ability to encode and reason about the associated spatial relational concepts. More specifically, we predict that children's performance on a midpoint task will be correlated with their knowledge of *middle* and *between*. To test this prediction, we gave children two tasks. First, children played a midpoint search task, in which they had to find a hidden object in the middle of two flags. No spatial terms were used during this task. Then children were given a language task to assess their production and comprehension of the words *middle* and *between* (as well as some other spatial terms). The relationship between children's performance on these two tasks was assessed. If, as we propose, specific spatial terms can facilitate encoding and reasoning about the corresponding spatial concepts, then we should find a relation between children's knowledge of *middle* and *between* and their performance on the midpoint task, even after controlling for age and other spatial word knowledge. By controlling for age, we ensure that any relationship found is not driven by other age-related factors that may influence both spatial vocabulary knowledge and relational encoding. Additionally, controlling for other spatial word knowledge allows us to distinguish relationships between spatial terms that describe midpoint, or key aspects of it—specifically, *middle* or *between*—and children's spatial vocabulary more generally.

In designing the task, we aimed to focus on children's *ability* to use the midpoint relation. In many empirical tests, it is unclear whether children's midpoint encoding behavior is a consequence of preference or ability. Previous investigations have mostly focused on children's spontaneous encoding strategies by measuring children's behavior on a single trial after ambiguous training. This leaves open the possibility that young children are able to encode midpoint but do not use it as a default strategy. (See Levelt, 2005, for a related discussion.) Children can switch spatial encoding strategies when their initial strategy is no longer viable (Marsh et al., 2011; Uttal et al., 2006). Thus, if children are able to encode the midpoint, and they receive feedback that a midpoint strategy is appropriate, they should be able to flexibly adopt a midpoint strategy following such feedback.

Alternatively, encoding and reasoning about the midpoint relation may pose a genuine challenge for young children. As noted earlier, non-human animals that have demonstrated midpoint encoding have typically needed extensive training. For these animals, it seems that a midpoint strategy is not simply dispreferred but difficult (e.g., Kamil & Jones, 1997; Kamil & Jones, 2000). Given that young children are known to find relational reasoning challenging (Gentner & Rattermann, 1991; Gentner & Toupin, 1986), it is important to discover whether children also find reasoning about the midpoint relation difficult. Therefore, we modified the standard midpoint task paradigm to include multiple test trials with feedback, giving children the opportunity to revise their encoding strategy. This allowed us to examine children's search behavior on the first expanded trial (as in prior studies), and to compare this with their performance after receiving feedback. If the midpoint encoding strategy is available, but simply non-preferred, then children should be able to switch to a midpoint search strategy given this feedback. However, if children genuinely find midpoint encoding challenging, then such feedback should not materially increase their success on the midpoint task.

2. Method

2.1. Participants

Fifteen 3-year-olds (30–41 months, M = 35.73), eighteen 4-year-olds (42–53 months, M = 46.94), and seventeen 5-year-olds (54–64 months, M = 58.65) recruited from the university's surrounding areas participated in this study. This age range was selected because children learn and continue to refine their knowledge of the words *middle* and *between* during this period (e.g., Johnston & Slobin, 1979; Hund, Bianchi, Winner, & Hesson-McInnis, 2017), and because previous work on children's midpoint encoding has also focused on this age range (e.g., Ankowski et al., 2012; MacDonald et al., 2004; Spetch & Parent, 2006; Uttal et al., 2006). An additional fourteen 3-year-olds, four 4-year-olds, and six 5-year-olds participated but were excluded due to failure to meet inclusion criteria (see Sections 2.2.1 and 3.1.1 for details). Parents provided written consent and children gave verbal assent to participate. All children participated in a laboratory on campus and received a small gift for their participation. All were monolingual English speakers.

2.2. Materials and procedure

Children participated in a single session, in a 6' × 10' testing room. The room had doors on three walls and cabinets along the fourth. A tripod and camera were placed in the corner to record the session, but no other objects were visible in the room, other than those used during the task itself. The Midpoint Task was administered first, followed by the Language Task.

2.2.1. Midpoint task

The purpose of the Midpoint Task was to assess children's spontaneous encoding of the midpoint relation, as well as their ability to use feedback to revise their encoding strategies. Children were first shown the location of an object hidden in a fixed configuration

Fig. 1. Midpoint Task Apparatus, including hiding box, landmarks (flags) and treasure chest.

(i.e., at the midpoint of two landmarks). To examine children's spontaneous encoding, the configuration was then expanded to a new distance to see whether children searched correctly at the midpoint or elsewhere, as previous research has done (e.g., MacDonald et al., 2004). To explore children's ability to revise their encoding strategies, we built upon this foundational task and gave children additional search trials with implicit feedback by always allowing children to discover the correct hiding location. The Midpoint Task typically took about 25 min to administer.

We used a small treasure chest with stickers inside as the target object and two flags – one red and one blue – as landmarks. The finding box measured 72" × 8" × 9" and was filled with Styrofoam packing peanuts (Fig. 1). The box remained in the same location for the duration of the task. Across a series of trials, the experimenter buried the treasure chest in the finding box and the child was asked to point to the treasure chest and then to dig to retrieve it. A ribbon placed on the back side of the finding box (visible only to the experimenter and replaced for each child) allowed the experimenter to mark the location of the child's searches. Children were not told the hiding rule, nor did the experimenter say "middle" or "between" at any point before or during administration of the Midpoint Task.

To ensure that children understood how to point during the Midpoint Task, children completed a pointing practice prior to the start of the task. During the pointing practice, the child watched the experimenter bury the treasure chest with no landmarks present. Children were instructed to point to the hidden treasure chest before digging to retrieve it. The experimenter provided feedback about the proper pointing and digging procedure. Children practiced pointing and digging on 2–4 trials, depending on how much practice was needed for the child to point and dig appropriately.

The Midpoint Task began with two training trials, in which the child watched the experimenter hide the treasure chest in the center of the two flags, placed 12 in. apart (Fig. 2). The experimenter then prompted the child to point to the buried treasure chest before allowing them to search. Children who were unable to point accurately on both training trials were excluded from further analyses (see Section 3.1.1 for more details).

After the training phase, children completed eight test trials. On test trials, the child closed their eyes while the experimenter counted to ten and buried the treasure chest. As a precaution against children being able to locate the treasure chest by audible cues, the experimenter always put both hands under the Styrofoam peanuts, in separate locations, while hiding the treasure chest. Throughout the study, the surface of the hiding material was smoothed to eradicate any visible cues to the hiding location before children were allowed to open their eyes. Children were prompted to point to the buried treasure chest before searching, and the location of the first point was used for data analyses. However, children were allowed to search until they found the treasure chest. This meant that children received feedback about the chest's location on each trial. Thus, children had multiple opportunities to revise their encoding strategies and had information about what encoding strategy was the optimal one (a midpoint strategy).

The locations and inter-landmark distances of the flags and the treasure chest varied across trials. The first test trial used the same inter-landmark distance as at training (12 in.) and served as a check that children had encoded the training configuration correctly. If children failed to point accurately on this non-expanded trial – for instance, if they perseverated and searched at the same location they had found the treasure chest on the previous training trial – they were given a second chance with the same configuration at a different location within the finding box. This second opportunity was provided to make allowances for the fact that children might not immediately recognize the connection between the training and test trials, and to therefore provide children with another chance to demonstrate their understanding of the original configuration. If children did not search correctly on the first test trial (on the first or second try), their data were excluded from further analyses (see Section 3.1.1 for more details). Children's performance on Trial 1 was coded as correct if they were accurate on either the first or second try.

The second test trial expanded the landmarks to 24 in. This was the first expanded trial and served as a measure of children's spontaneous encoding strategies. In this expanded array, it was possible to distinguish children who encoded location relationally – which should lead a child to search correctly at the midpoint – from other strategies. For instance a child using a vector strategy would search not at the midpoint but at the distance (i.e., 6 in.) and direction seen at training.

The remaining six test trials alternated between non-expanded (i.e., 12-inch) and expanded (24- or 36-inch) distances (Fig. 2). Alternating expanded trials with non-expanded trials ensured that children's encoding strategies on expanded trials could be identified, whether they were encoding location based on the training trials or the previous trial.

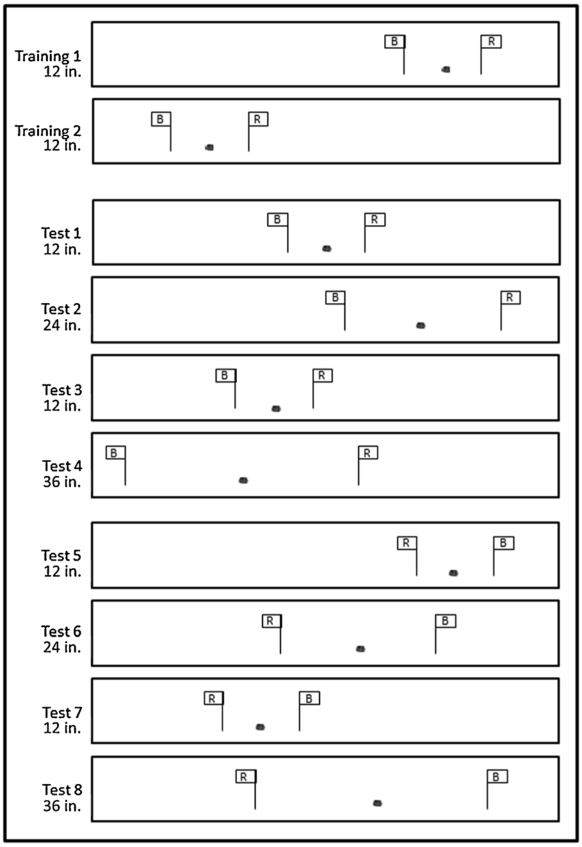

Fig. 2. Midpoint Task training and test trials.

2.2.2. Language task

The Language Task immediately followed the Midpoint Task. The chief purpose of this task was to assess children's production and comprehension of the spatial prepositions *middle* and *between*. We also tested three other spatial prepositions (*in*, *on*, *under*) that are typically acquired during early preschool (Johnston, 1988; Tomasello, 1987). On each trial, children either described a spatial relation between a target object and one or more ground object(s) (in the production trials) or chose the appropriate configuration to fit the experimenter's description (in the comprehension trials). The Language Task typically took about 15 min to administer.

A small plastic cow and pig served as target objects. The ground objects were unpainted wooden figurines of artifacts, vehicles, and furniture (Fig. 3). To ensure that children knew the names of the objects and object categories that would be used during the task, children were asked to name all the objects before beginning the task. Children were told the correct labels for any items they could not name, and if a child gave an appropriate alternative label (e.g., "cradle" instead of "crib"), the experimenter adopted that term. After the child named all the individual items, the experimenter presented all of the vehicles together and told the child that sometimes these are called *vehicles*; the same was done for the *furniture* items.

The Language Task was administered in three sections: production, placement comprehension, and forced-choice comprehension. Trials were administered in a fixed order, with two examples[3] of each spatial relationship in each section, resulting in eight production, ten placement comprehension, and ten forced-choice comprehension trials (a total of 28). See Appendix A for more details.

On production trials, children were shown arrays of two objects and a small plastic cow and asked, "Where's the cow?" (Fig. 3a). In trials testing *middle/between*, the cow was always placed at the midpoint between two objects, and both *between* and *middle* were considered correct responses for these trials. On placement comprehension trials, children were also shown arrays with two objects but were asked to put a small plastic pig in the appropriate relation (e.g., "Put the pig between the cars.") (Fig. 3b). For *between* trials, placement anywhere between the ground objects was considered correct. For *middle* trials, placement in the center third of the space between ground objects was considered correct. Finally, on forced-choice comprehension trials, children were shown two arrays, identical except that one showed a pig in a certain relation and the other showed a cow in a different relation (e.g., a pig between the vehicles and a cow next to a vehicle). The children were then asked, for example, "Is the pig or the cow between the vehicles?" For *between* trials, the arrays contrasted a figure next to and between the ground objects (correct) with a figure next to but not between

[3] In the production section, children could have used either "middle" or "between" to accurately describe the arrays depicting those relations; thus, we did not assess these terms separately in this section.

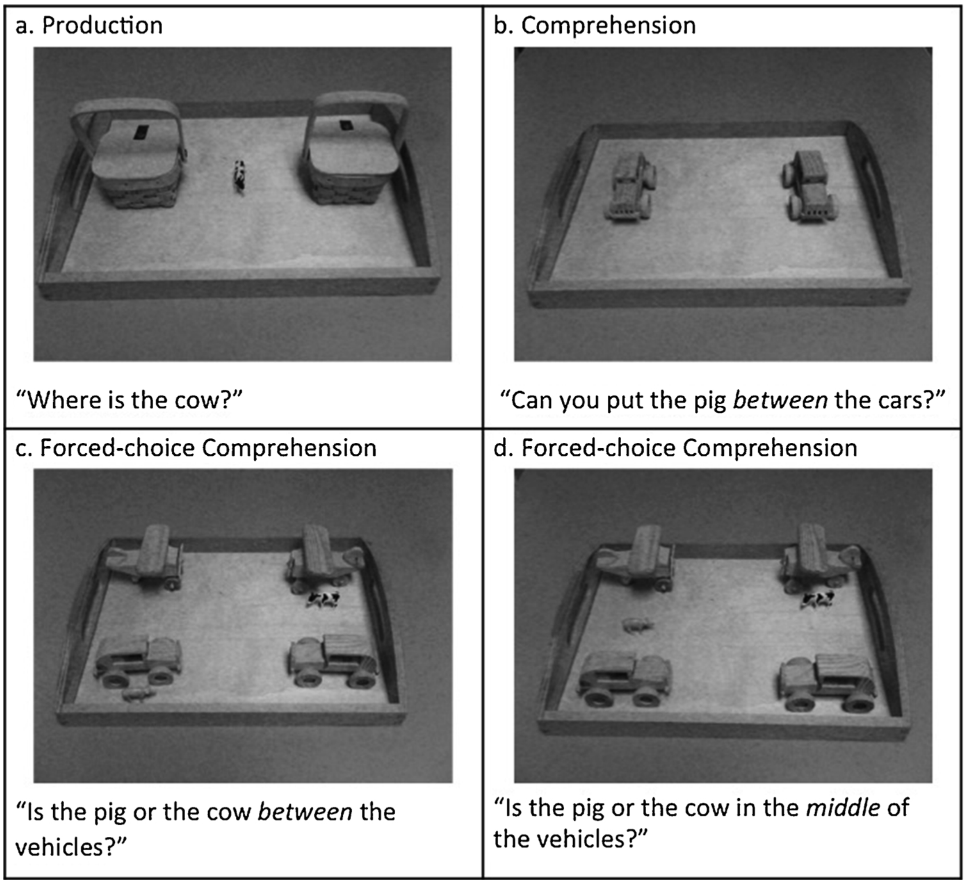

Fig. 3. Example *middle/between* arrays from the Language Task.

the ground objects (Fig. 3c). For *middle* trials, the two arrays contrasted a figure at the center of the two ground objects (correct) with a figure between, but not at the center, of the two ground objects (Fig. 3d).[4]

3. Results

Before describing the key analyses relating Midpoint Task performance to Language Task performance, we first describe the data from each task individually.

3.1. Midpoint task

3.1.1. Coding and exclusion criteria

During the task, the location of each child's points were marked by the experimenter on a length of ribbon and measured to the nearest half inch. Each location was then coded as correct or incorrect. To allow for some imprecision in children's pointing, responses within one-and-a-half inches on either side of the correct location were coded as correct on all 12-inch trials. A slightly larger margin was used for expanded trials: on 24-inch trials, children had to point within two-inches on either side of the treasure chest, and on 36-inch trials, they had to point within two-and-a-half-inches. These margins accounted for the fact that it becomes more difficult to perceive the midpoint as the inter-landmark distance increases, while still allowing enough precision to distinguish between different encoding strategies (see below).

Children who did not point within this margin on both training trials were excluded from further analysis (7 three-year-olds, 3 four-year-olds, 1 five-year-old), as they either did not understand the task instructions or could not accurately indicate the location of the treasure chest by pointing even when it was known. Children who were not able to correctly point to the treasure chest on the first, non-expanded test trial (either on the first or second try) were also excluded from further analysis (an additional 6 three-year-olds, 1 four-year-old, 4 five-year-olds), as they may not have been able to remember the training configurations accurately. Because

[4] We note that our task assesses children's understanding of *middle* as it refers to the midpoint, which is only one of several senses of the term. However, this sense is the one predicted to provide the best support for midpoint encoding in the search task.

Table 1
Response types on first (Test Trial 2) and last (Test Trial 8) expanded trial.

	Trial	3-year-olds	4-year-olds	5-year-olds	Total
Correct	First	4 (27%)	9 (50%)	15 (88%)	28 (56%)
	Last	1 (9%)	10 (59%)	12 (75%)	23 (52%)
Between flag errors	First	7 (47%)	8 (44%)	2 (12%)	17 (34%)
	Last	10 (91%)	7 (41%)	4 (25%)	21 (48%)
Outside flag errors	First	4 (27%)	1 (6%)	0 (0%)	5 (10%)
	Last	0 (0%)	0 (0%)	0 (0%)	0 (0%)
n	First	15	18	17	50
	Last	11	17	16	44

this trial was used as an inclusion criterion, it was not included in further analyses of children's test performance.

Incorrect responses were further coded into four error types, corresponding to particular encoding strategies, as discussed earlier: *Beacon*, *Vector*, *Perseverative*, and *Other*. Because these codes are not central to our research questions, we do not discuss these further in the Results section. However, for comparability with prior work on midpoint encoding, we report these in Appendix B.

To determine whether children at least understood that the treasure chest was to be found between the flags, even if they were unable to precisely encode or indicate the midpoint, we also coded errors more generally as *Between* or *Outside* the flags. This coding system was largely orthogonal to the first coding system (e.g., a perseverative error might occur between or outside the flags).

Not all children completed every trial. The experimenter ended the Midpoint Task early if children asked to stop playing, or if they became unresponsive during the task and confirmed that they wanted to stop playing when asked. Occasionally, children also failed to provide a recordable response for a trial—for example by digging before pointing. Trials on which children did not point were treated as missing data. Because of our interest in both children's spontaneous strategy use and their ability to revise their strategy after feedback, children who did not complete at least one trial after receiving disambiguating feedback from the first expanded trial were excluded from further analyses (one additional 3-year-old and one 5-year-old). Of the final sample, 88% of children had data for all 7 trials (3-year-olds: 73%; 4-year-olds: 94%; 5-year-olds: 94%). Children completed an average of 6.72 trials (3-year-olds: $M = 6.47$; 4-year-olds: $M = 6.83$; 5-year-olds: $M = 6.82$).

3.1.2. Default strategy use

To assess children's initial use of a midpoint strategy prior to feedback, we examined children's responses on the first expanded trial (Test Trial 2). Table 1 shows the number of children in each age group who searched correctly on the first expanded trial, as well as whether their errors were between or outside the flags. As expected, the number of errors diminished with age, $\chi^2(2, N = 50) = 12.67, p < .01$. Pairwise comparisons revealed that 5-year-olds made significantly fewer errors on this trial than 3-year-olds, $p < .01$, or 4-year-olds, $p < .05$, whose performance did not differ significantly.

Strikingly, the youngest children in our study often searched outside the landmarks; 27% of the 3-year-olds (and one 4-year-old) made Outside Flag errors. Although these children were able to replicate the configuration they had seen at training on the first, non-expanded test trial, they did not abstract away from this fixed arrangement, even as far as to encode location qualitatively as between the flags.

3.1.3. Strategy revision following feedback

To assess whether children's midpoint search performance improved after trials with feedback, we report the response types on the last expanded trial (Test Trial 8; Table 1). The proportion of children who responded correctly on the last trial was not significantly different from that on the first expanded trial (exact McNemar's test, $p = .77$). That is, we saw no evidence that children possessed a midpoint hypothesis that was simply dispreferred.

One notable point of improvement, however, was in the nature of children's errors: on the first expanded trial, 27% of the youngest children searched outside the flags, suggesting a failure to encode even the qualitative *between* aspect of midpoint. In contrast, on the last expanded trial, these children made no such errors – all their searches took place between the two flags. Thus, although feedback on each trial did not lead children to adopt a precise midpoint strategy, it did seem to allow the youngest children to gain the insight that the hiding location was between the flags.

3.1.4. Overall performance

To assess children's overall performance, we analyzed children's performance over all seven test trials. Because some children did not have data for every trial, children's responses were converted to proportions out of their total number of trials. The mean proportions of correct responses and between-/outside-flag errors are given in Table 2. Children significantly improved in their overall midpoint search accuracy with age (in months), $r = .65, p < .001$. While most of children's errors fell between the landmarks, the youngest age group made a surprisingly high proportion of searches outside the landmarks, with 19% of their responses landing outside the flags, again reflecting a pattern seen on the first expanded trial. Children showed a significant decrease in Outside responding with age (in months), $r = -.61, p < .001$.

Table 2
Mean Proportion of Correct Responses and Between- and Outside-flag Errors on Midpoint Task.

	3-year-olds *M* (*SD*)	4-year-olds *M* (*SD*)	5-year-olds *M* (*SD*)	Total *M* (*SD*)
Out of Total Responses[Ω]	*n* = 15	*n* = 18	*n* = 17	*n* = 50
Correct	0.37 (0.24)[b,c]	0.65 (0.30)[a]	0.83 (0.16)[a]	0.63 (0.30)
Between flag errors	0.44 (0.15)[c]	0.30 (0.26)	0.17 (0.16)[a]	0.30 (0.22)
Outside flag errors	0.19 (0.16)[b,c]	0.06 (0.11)[a]	0.00 (0.00)[a]	0.08 (0.13)
Out of Total Errors	*n* = 15	*n* = 14	*n* = 12	*n* = 41
Between flag errors	0.74 (0.22)	0.87 (0.19)	1.00 (0.00)	0.86 (0.20)
Outside flag errors	0.26 (0.22)	0.13 (0.19)	0.00 (0.00)	0.14 (0.20)

Notes: These means are for all seven test trials, including the first expanded trial.
[Ω] Bonferroni corrected pairwise comparisons, $p < .05$.
[a] Different from 3-year-olds.
[b] Different from 4-year-olds.
[c] Different from 5-year-olds.

3.2. Language task

As in the Midpoint Task, not all children finished every trial. One 5-year-old chose not to participate in the Language Task. The remaining children sometimes refused to answer a given question. If a child asked to stop or grew unresponsive on a particular section of the Language Task, the experimenter moved on to the next section. Overall, 80% of the children completed all 28 trials of the Language Task (3-year-olds: 60%; 4-year-olds: 83%; 5-year-olds: 94%). On average, children completed 25.20 trials (3-year-olds: $M = 22.13$, 4-year-olds: $M = 25.67$, 5-year-olds: $M = 27.41$). Proportion correct was computed out of the total completed for each measure.

Middle was our primary interest in the Language Task, as a label for the midpoint relation. We also included *between* as a term that captures a critical aspect of the midpoint relation and which may also confer some benefit to children's midpoint encoding. However, because of the small number of each type of trial on the Language Task for each spatial term, even a small proportion of missing trials reduced our confidence in analyzing individual terms separately. Therefore, we conducted our main analyses on the spatial terms collapsed into two sets: those that capture important information for midpoint encoding (*middle* and *between*) and those that do not (*on*, *in*, and *under*). We report the production and comprehension rates of *middle* and *between* separately in Appendix C.

However, here we highlight one suggestive pattern from the *middle* and *between* forced choice comprehension trials. Of the children who completed both *between* trials, children who got both correct were categorized as *between*-knowers. Likewise, of the children who completed both *middle* trials, children who got both correct were categorized as *middle*-knowers (Table 5). No 3-year-olds were *middle*- or *between*-knowers (0 out of 9 in both cases). However, by 5-years-old most children were *between*-knowers (12 out of 16), but most were still not *middle*-knowers (6 out of 16). These patterns are consistent with the possibility that children initially understand *middle* to mean something akin to *between.* Incorporating centrality into the meaning of *middle* may take longer.

Table 3
Proportion Correct on *Middle/Between* and *On/In/Under.*

Spatial terms	Subtest (# trials)	3-year-olds *M* (*SD*)	4-year-olds *M* (*SD*)	5-year-olds *M* (*SD*)	Total *M* (*SD*)
middle/between	Production (2)	0.25 (0.45) *n* = 12	0.58 (0.49) *n* = 18	0.85 (0.34) *n* = 17	0.60 (0.48) *n* = 47
	Comprehension: Placement (4)	0.52 (0.40) *n* = 14	0.82 (0.33) *n* = 17	0.90 (0.18) *n* = 17	0.76 (0.34) *n* = 48
	Comprehension: Forced Choice (4)	0.20 (0.16) *n* = 10	0.52 (0.36) *n* = 16	0.61 (0.36) *n* = 16	0.48 (0.36) *n* = 42
	Total (10)	0.37[b,c] (0.27) *n* = 14	0.63[a] (0.28) *n* = 18	0.77[a] (0.19) *n* = 17	0.60 (0.29) *n* = 49
on/in/under	Production (6)	0.74 (0.27) *n* = 15	0.92 (0.12) *n* = 18	0.97 (0.07) *n* = 17	0.88 (0.19) *n* = 50
	Comprehension: Placement (6)	0.87 (0.18) *n* = 14	0.96 (0.08) *n* = 17	0.98 (0.08) *n* = 17	0.94 (0.12) *n* = 48
	Comprehension: Forced Choice (6)	0.88 (0.25) *n* = 10	0.96 (0.10) *n* = 16	1.00 (0.00) *n* = 16	0.96 (0.14) *n* = 42
	Total (18)	0.83[b,c] (0.16) *n* = 15	0.93[a] (0.09) *n* = 18	0.98[a] (0.03) *n* = 17	0.92 (0.12) *n* = 50

Bonferroni corrected pairwise comparisons, $p < .05$.
[a] Different from 3-year-olds.
[b] Different from 4-year-olds.
[c] Different from 5-year-olds.

Table 4
Regression models predicting Midpoint Task accuracy with age in months and Language Task accuracy (*middle/between* and *on/in/under*).

	R^2	ΔR^2	Age in months β (SE)	*middle/ between* β (SE)	*on/in/ under* β (SE)
Model 1	0.435***	–	0.021*** (0.003)	–	–
Model 2	0.492***	0.057*	0.016*** (0.004)	0.300* (0.132)	–
Model 3	0.492***	0.000	0.016** (0.004)	0.300* (0.137)	0.002 (0.339)

ˆ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.
Note: ΔR^2 is the change in R^2 from the preceding model, e.g., from Model 1 to Model 2 or from Model 2 to Model 3.

Table 3 shows the mean proportion correct for the various subsections of the Language Task on the two sets of spatial terms (*middle/between* and *on/in/under*). One 3-year-old did not complete any *middle/between* trials, and so their data were excluded from further analysis, leaving a total of 49 participants. Both *middle/between*, $r = 0.56$, $p < .001$, and *on/in/under* knowledge, $r = 0.55$, $p < .001$, was significantly related to age (in months). However, after controlling for their relationships with age, performance on *middle/between* and *on/in/under* trials were not correlated, suggesting that knowledge of *middle* and *between* was independent from *on/in/under*.

3.3. *Word knowledge and midpoint encoding*

3.3.1. *Middle/between and On/in/under*

To understand how age and knowledge of each set of spatial terms contributed to children's ability to encode the midpoint, proportion correct on the Midpoint Task was predicted using a step-wise linear regression analysis with age in months entered into Model 1 (Age only), proportion correct on *middle/between* trials from the Language Task added into Model 2 (Age + *middle/between*), and proportion correct on *on/in/under* trials from the Language Task added into Model 3 (Age + *middle/between* + *on/in/under*). The three models are summarized in Table 4.

All three models significantly predicted Midpoint Task performance (Age alone: $F(1,47) = 36.25$, $p < .001$; Age + *middle/between*: $F(2,46) = 22.32$, $p < .001$; Age + *middle/between* + *on/in/under*: $F(3,45) = 14.55$, $p < .001$). Age alone accounted for 43.5% of the variance on the task. Importantly, adding *middle/between* accuracy significantly improved the fit of the model, $F(1,46) = 5.17$, $p < .05$, and accounted for 49.2% of the variance. Both age and *middle/between* knowledge were significant individual predictors of Midpoint Task performance. However, adding *on/in/under* accuracy did not improve the fit of the model, $F(1,45) = 0.00$, $p = .99$. In the final model including age, *middle/between*, and *on/in/under* performance, only age and *middle/between* were maintained as individual predictors of Midpoint Task performance.

3.3.2. *Middle and between*

Although we did not conduct formal analyses of the relationship between Midpoint Task performance and *middle* and *between* separately, we report the proportion of correct responses and responses that fell outside the flags for *between*-knowers (children who got both *between* forced-choice trials correct) and *middle*-knowers (children who got both *middle* forced-choice trials correct) in Table 5. (It is important to note that there is overlap in these groups – some children may be classified as both *middle*- and *between*-knowers.)

Overall, *between*-knowers and *middle*-knowers were both more accurate than non-knowers on the Midpoint Task. They also made fewer outside-flag responses. In this sample, the group of *middle*-knowers seems to have been more accurate and made fewer outside responses than the group of *between*-knowers. It is possible that *middle* provides better support for midpoint encoding, in that it more aptly describes the midpoint relation than *between* does.

4. General discussion

The theoretical framework for this research is that relational language can support the acquisition and use of relational representations (e.g., Gentner, 2003, 2010, 2016). Accordingly, we hypothesize that knowledge of the spatial terms *middle* and *between* supports children's ability to reason about the midpoint relation. Our findings are consistent with this hypothesis—though as discussed later, other interpretations are possible. A further goal was to discover whether encoding the midpoint is genuinely difficult for children, or whether it is merely a dispreferred strategy. Our findings support the first view.

4.1. *Spatial language supports spatial thinking*

We found that children's performance on a nonlinguistic midpoint search task was predicted by their understanding of *middle* and

Table 5
Mean Proportion Correct and Outside-flag Responses on Midpoint Task for *Between*-knowers[a] and *Middle*-knowers[b].

Midpoint Task Responses	Forced Choice Comprehension	3-year-olds *M* (*SD*)	4-year-olds *M* (*SD*)	5-year-olds *M* (*SD*)	Total *M* (*SD*)
Between					
Correct responses	Knowers	– $n = 0$	0.71 (0.29) $n = 6$	0.88 (0.13) $n = 12$	0.83 (0.21) $n = 18$
	Non-knowers	0.35 (0.27) $n = 9$	0.61 (0.33) $n = 10$	0.75 (0.14) $n = 4$	0.53 (0.32) $n = 23$
Outside-flag responses	Knowers	– $n = 0$	0.05 (0.07) $n = 6$	0.00 (0.00) $n = 12$	0.02 (0.05) $n = 18$
	Non-knowers	0.19 (0.17) $n = 9$	0.06 (0.14) $n = 10$	0.00 (0.00) $n = 4$	0.10 (0.16) $n = 23$
Middle					
Correct responses	Knowers	– $n = 0$	0.86 (0.18) $n = 6$	0.88 (0.11) $n = 6$	0.87 (0.14) $n = 12$
	Non-knowers	0.35 (0.27) $n = 9$	0.53 (0.31) $n = 10$	0.83 (0.16) $n = 10$	0.58 (0.32) $n = 29$
Outside-flag responses	Knowers	– $n = 0$	0.00 (0.00) $n = 6$	0.00 (0.00) $n = 6$	0.00 (0.00) $n = 12$
	Non-knowers	0.19 (0.17) $n = 9$	0.09 (0.14) $n = 10$	0.00 (0.00) $n = 10$	0.09 (0.14) $n = 29$

[a] *Between*-knowers were children who were correct on both of the forced-choice *between* trials. Non-knowers completed both trials but had only one or none correct.

[b] *Middle*-knowers were children who were correct on both of the forced-choice *middle* trials. Non-knowers completed both trials but had only one or none correct.

between on a subsequent language-assessment task. This connection was specific: children's facility with the midpoint relation was predicted by knowledge of *middle* and *between,* but not by knowledge of other spatial terms (*in*, *on* and *under*). Further, although children's performance improved with age, their knowledge of relevant spatial terms predicted performance beyond what was accounted for by age. Thus, it appears that children with robust knowledge of words like *middle* and *between* are better able to encode and use the midpoint relation in a nonlinguistic task.

Our findings are consistent with prior evidence suggesting that knowledge of the word *middle* predicts children's ability to encode and use the midpoint relation (Ankowski et al., 2012; MacDonald et al., 2004). More generally, our findings fit with other research showing connections between children's spatial language and their spatial reasoning (Balcomb et al., 2011; Gentner et al., 2013; Hermer-Vazquez et al., 2001; Loewenstein & Gentner, 2005; Miller et al., 2016; Pruden et al., 2011; Pyers, Shusterman, Senghas, Spelke, & Emmorey, 2010). While the correlational nature of our design does not permit conclusions about causal relationships between children's word knowledge and midpoint encoding, our findings are consistent with a view of spatial language as a tool for spatial thinking (Gentner, 2010, 2016; Gentner & Christie, 2012). In this view, learning and using *middle* and *between* fosters children's ability to encode the midpoint relation. Thus, a child who has knowledge of the word *middle* will be more likely to encode a location as *in the middle*, even when the particulars of the instances (e.g., size of the array) vary, and this facilitates generalization of the midpoint relation across instances.

Evidence that spatial language supports spatial thinking comes from a variety of sources. For example, Loewenstein and Gentner (2005) found that preschool children given a difficult spatial mapping task performed better when the task was introduced using the terms (*top*, *middle* and *bottom*) that highlighted the spatial relations that were important for the task. The label advantage persisted over a two-day delay, without reinstating the terms, suggesting that the labels acted to promote a more delineated relational representation, rather than simply as a temporary crutch. Hearing labels for spatial relations has also been found to support children's learning of spatial relational categories (Casasola, 2005; Casasola & Bhagwat, 2007; Casasola et al., 2009).

Other evidence comes from comparing languages with different patterns of spatial language. Haun et al. (2006) compared Dutch speakers, whose language (like English) primarily uses an egocentric frame of spatial reference, with speakers of Hai||om (a Khoisan language spoken in Namibia), which primarily uses a geocentric frame. Participants saw an array of five identical cups on a table and watched as an object was hidden in one of the cups. They then moved around to a second table, where they saw an identical array of cups, but from the opposite direction. Their task was to find a second hidden object, based on the location of the first hidden object. The objects were hidden according to fixed rules following either an egocentric or geocentric encoding. People received feedback on each trial, and both groups improved over trials. The results showed that people learned fastest when the mapping rule fit their spatial language: Dutch speakers performed best when left-right position was maintained (consistent with their egocentric linguistic pattern), and Hai||om speakers performed best when north-south position was maintained (consistent with their geocentric linguistic pattern). Thus, people's habitual language led to greater facility with the associated encoding strategy.

Further evidence comes from comparing speakers who vary in their knowledge of spatial language. For example, different cohorts of speakers of Nicaraguan Sign Language (NSL—a relatively new and rapidly developing sign language; Senghas, 2003) vary in how consistently they mark *left* and *right* relations, and these differences predict success on the reorientation task (Pyers et al., 2010). In the same vein, Gentner et al. (2013) compared 6-year-old children who lack spatial language—deaf children who had not been taught a standard sign language, and who instead used their own invented homesign systems (see Goldin-Meadow, 1993)—with a matched group of hearing children. When both groups were given a nonlinguistic version of the Loewenstein and Gentner (2005) spatial mapping task, the hearing children performed very well, whereas the homesigners performed at chance. In a subsequent language assessment task, the hearing children produced a full range of spatial language; in contrast, the homesigners produced hardly any gestures for spatial relations, despite fluently producing gestures for objects and actions. These results suggest a strong connection between possessing spatial language and performing well on spatial tasks. We view this connection as representative of a larger pattern in which relational language supports learning and reasoning about relational information (Gentner, 2003, 2010, 2016).

How might this connection come about? According to Gentner's (2003, 2010, 2016; Gentner & Christie, 2012; Gentner & Loewenstein, 2002) account, there are at least four mechanisms by which language acquisition fosters children's understanding of the corresponding concepts. First, during language learning, hearing a linguistic label applied to different exemplars fosters comparison and abstraction across the exemplars; language thus serves to jump-start the learning of relational concepts that might otherwise be learned much later (if at all). Second, naming promotes reification: a linguistic label invites storing a unified representation of the relational pattern, making it more accessible for future use (Jamrozik and Gentner, 2013, 2015; see also Lupyan et al., 2007; Son et al., 2012). Third, naming promotes uniform relational encoding: Habitual use of a given term makes it more likely that the relational constellation will be encoded in the same manner across contexts. This is especially important for relational concepts, because evidence suggests that relations are encoded in a more context-specific way than objects (Bassok, Chase, & Martin, 1998; Forbus, Gentner, & Law, 1995; Gentner, Loewenstein, Thompson, & Forbus, 2009; Kersten & Earles, 2004). Finally, a fourth possible influence of language is Quinean bootstrapping (Carey, 2009): hearing a new term used in context may invite children to form of placeholder for a meaning that they are still learning.

Because the midpoint relation is a complex spatial relation requiring the integration of multiple relationships, we think it likely that some or all of these mechanisms contribute to children's learning. Children's word knowledge – including, but not limited to relational terms – may also support children's task performance in other ways, for instance by providing a means for self-regulation and the creation and maintenance of task goals (e.g., Jacques & Zelazo, 2005; Kirkham, Cruess, & Diamond, 2003; Miller & Marcovitch, 2011).

Importantly, however, we must acknowledge that our results, while suggestive, do not allow us to conclude that language had a causal effect on children's spatial representations in our task. Perhaps it is the reverse: that the ability to encode the midpoint relation is what allows children to learn the corresponding terms. A third possibility is that some other general factor, such as high spatial ability, or a supportive early environment, leads to superior performance in both our tasks. Additional research will be needed to tease apart these causal relationships.

4.2. The midpoint relation is challenging

This study was the first to disentangle children's *ability* to encode the midpoint relation from their default encoding pattern. In prior research, children were trained on fixed-distance arrays and then were given a single expanded test trial (e.g., MacDonald et al., 2004; Spetch & Parent, 2006; Uttal et al., 2006). While this provides information about children's spontaneous strategy, it does not tell us what they would do if given further information about the optimal encoding strategy. Thus, young children's failure to use a midpoint strategy may have been due simply to a preference for other encoding strategies, rather than to finding the midpoint relation challenging. To put it another way, perhaps children have a midpoint encoding hypothesis, but other encoding hypotheses are higher in a priori likelihood.

To examine this question, we used a training method, as in Haun et al.'s (2006) study. We first assessed children's spontaneous performance on the first expanded trial. Then we gave them a series of trials with feedback as to the correct answer. Our findings show that young children genuinely find encoding and reasoning about the midpoint relation challenging. Performance on the last expanded trial (after 6 trials with feedback) was no higher than performance on the first expanded trial. This conclusion is also consistent with Ankowski et al. (2012) finding that although hearing the hiding location described as "in the middle" marginally increased midpoint responding, it did not allow all children (especially the youngest) to immediately adopt a midpoint strategy, as might have been expected if the verbal cue allowed children to simply select the appropriate encoding from among various accessible options.

4.3. Additional insights into midpoint development

4.3.1. Qualitative precedes quantitative: between before middle

Although children did not show gains in correct midpoint responses, we did see evidence of some learning in the youngest group. Initially, 27% of the youngest age group searched outside the landmarks. Searches outside the landmarks reveal a failure to take both landmarks into consideration when encoding location. These 3-year-olds were not merely lacking precision – they had not even encoded location qualitatively as *between* the flags. One intriguing question that arises from these findings is whether the midpoint relation and the between relation are initially conflated, and only gradually differentiated.

In support of this conjecture, we point to three of our findings. First, the overall proportion of children's responses that fell outside the flags (rather than between the flags) steadily declined across age in our study (from 19% for 3-year-olds to 0% for 5-year-olds). That is, children seemed to increasingly appreciate that the location was between the flags, even if they did not search precisely at the midpoint. Likewise, Spetch and Parent (2006) found that while both younger and older children often searched at locations adjacent to the landmarks (rather than in the middle), younger children were much more likely to do so outside the landmarks than older children, who primarily searched at the adjacent location between landmarks. A second point is that whereas 27% of 3-year-olds (and one 4-year-old) searched outside the flags on the first expanded trial, no children searched outside the flags on the last expanded trial. Thus, repeated feedback helped children to gain the important insight that the location was between the flags, even though they did not improve in their midpoint encoding. This suggests that understanding the qualitative between relation may be a first step in acquiring the midpoint relation.

Finally, children's performance on the Language Task is also suggestive of a qualitative-first pattern of acquisition. Children's comprehension of *between* outpaced their comprehension of *middle*. There were fewer children by 5-years-old who could discriminate *middle* from *between* (6/16, as measured by the forced-choice *middle* trials) than who could discriminate *between* from *next to* (12/16, as measured by the forced-choice *between* trials), suggesting that this qualitative relationship is learned earlier. Thus, we suggest that children may understand *middle* to mean something like *between* initially, and only later incorporate more precision. Development of a complete, adult-like understanding of spatial relational terms, especially complex ones, can be protracted, with some aspects of the meaning being acquired before others (e.g., Durkin, 1981; Weber, Miller, & Ou, 2018; see also Clark, 1973; Gentner, 1978; Tillman & Barner, 2015; Wagner, Jergens, & Barner, 2018). Adults have multiple senses of the word *middle*, with one that refers to the midpoint. Young children's experience with the word *middle* may not initially lead them to this more specific sense of its meaning. However, understanding that *middle* can refer to the midpoint should support children's encoding of the relation.

4.3.2. Many factors influence midpoint encoding

Although the main goal of this work was to explore the role of spatial language in the development of midpoint encoding, there are many factors that may influence children's success. Studies exploring the development of midpoint encoding – including this one – have varied along a number of dimensions, especially in regards to features of the search spaces used. For example, some studies have used 2-landmark arrays (the present study; Spetch & Parent, 2006; Uttal et al., 2006), while others have used 4-landmark arrays (Ankowski et al., 2012; MacDonald et al., 2004). Some studies have used spaces with discrete locations (e.g., cups: MacDonald et al., 2004, Exp. 1b; or boxes: Spetch & Parent, 2006), whereas others have used continuous search spaces (boxes filled with packing material, sand, or confetti: the present study; Ankowski et al., 2012; MacDonald et al., 2004, Exp. 2; a field: Uttal et al., 2006). And all but Uttal and colleagues' paradigm have used small-scale search spaces.

In addition, these studies have reported a wide range of ages at which children successfully demonstrated midpoint encoding. Although direct comparison between individual studies is precluded by their many differences, the patterns of results suggest some features that may affect the fluency of children's midpoint encoding. For example, children may be less likely to use a relational coding strategy in a subdivided, discrete space than a continuous space (Jeong, Levine, & Huttenlocher, 2007). MacDonald et al. (2004) found that older (5- to 9-year-old) children were less likely to search at the midpoint of four landmarks in a discrete space than were younger (3- to 4-year-old) children in a continuous space, though they did not directly compare same-age children on the two apparatuses. Likewise, Spetch and Parent (2006) found that their 3- to 5-year-olds had trouble even meeting the training criterion on the initial (non-expanded) trials when trained on two landmarks in a row of boxes; in contrast, children in this age range were mostly successful on these trials in the present study, which used a two-landmark configuration in a continuous space. Outside-configuration errors were also reported more frequently on studies testing children in discrete rather than continuous spaces (MacDonald et al., 2004; Spetch & Parent, 2006; Uttal et al., 2006). However, the continuous/discrete distinction does not always lead to performance differences: Marsh et al. (2011) found qualitatively similar patterns of searching on an iPad display made to look discrete (through the use of grid lines) versus continuous (no grid lines) with children spanning a broad range of ages.

Some reports have suggested that the number of landmarks in a configuration may influence midpoint encoding, though it is not clear whether more or fewer are expected to be beneficial (Potì et al., 2010; Sturz & Katz, 2009; Uttal et al., 2006). Ankowski and colleagues reported roughly comparable rates of midpoint searching on the first expanded trial in a four-landmark array as we do in the present study using two landmarks (both in continuous spaces). Another factor may be scale. Uttal and colleagues, using a large-scale search task (also with two landmarks in a continuous space), found much higher midpoint searching than either Ankowski et al. or the present study—suggesting that larger scale may support children's midpoint search performance. This is consistent with findings from other spatial reasoning tasks, which showed that larger spaces led to more successful landmark use in younger children (Learmonth, Nadel, & Newcombe, 2002).

These results highlight the need for future work to more systematically explore how features of the spatial array and environment encourage (or discourage) midpoint encoding, and how they influence spatial encoding strategies more generally. For instance, configurations that can be united into a systematic, meaningful unit (e.g., a familiar shape) may facilitate children's spatial representations (Uttal, Gregg, Tan, Chamberlin, & Sines, 2001).

Beyond the features of the search spaces, concurrently developing skills and processes are likely to influence midpoint encoding. For example, spatial scaling – the ability to translate between spatial arrays or representations of different (but proportional) sizes – is an integral spatial ability that shows considerable improvement between the ages of 3- and 5-years-old (Frick & Newcombe, 2012). In general, children also become more precise at encoding and remembering location (e.g., Huttenlocher, Newcombe, & Sandberg, 1994; Spencer & Hund, 2002; Spencer & Hund, 2003). Midpoint and related concepts (e.g., half) may play an important role in the development of children's spatial and mathematical reasoning more generally (e.g., Spinillo & Bryant, 1991), so understanding how language and other factors influence its acquisition may have broad implications.

4.4. Conclusion

There are two main findings. First, our results show that encoding and using the midpoint relation is difficult for young children.

Whereas human adults fluently and readily encode spatial location in terms of the midpoint, young human children (like nonhuman animals) do not—even with repeated feedback. Second, proficiency in using the midpoint relation is predicted by knowledge of the corresponding linguistic terms, *middle* and *between*—consistent with the idea that this challenging concept is acquired in part through learning spatial language. Human superiority in using the midpoint relation—and other challenging relations—may stem in part from possessing language that supports relational representation and reasoning.

Declarations of interest

None.

Author's note

An earlier and less complete report of this study was published in a CogSci proceedings paper.

Acknowledgements

This project was funded by SILC (the Spatial Intelligence and Learning Center), an NSF Science of Learning Center, grant SBE-0541957. The NSF had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication. The authors would like to thank Jen Hellige Palmer, the Project on Children's Thinking, and the parents and children who participated. We also thank David Uttal and the Cognition and Language Lab for helpful feedback.

Appendix A

Language Task Trials

Production: 'Where is the cow?'	
Target Word [Ground]	Objects in Display
in [crib]	Crib, plane
on [plane]	Table, plane
under [table]	Basket, table
middle/between [baskets]	Basket, basket[a]
under [car]	Cabinet, car
in [cabinet]	Cabinet, plane
middle/between [planes]	Plane, plane[a]
on [basket]	Basket, car
Comprehension Placement: 'Can you put the pig *X* the [ground]?'	
Target Word [Ground]	Objects in Display
under [crib]	Car, crib
in [car]	Table, car
between [tables]	Table, table
on [cabinet]	Cabinet, basket
middle [cribs]	Crib, crib
in [basket]	Car, basket
between [cars]	Car, car
under [airplane]	Plane, table
middle [cabinets]	Cabinet, cabinet
on [table]	Table, cabinet
Comprehension Forced Choice: 'Is the pig or the cow *X* the [ground]?'	
Target Word [Ground]	Objects in Display
in [car]	2 cars (*in* vs. *under*)
under [table]	2 tables (*on* vs. *under*)
between [furniture]	2 sets: crib, cabinet (*next to* vs. *between*) [b,d]
on [airplane]	2 airplanes (*under* vs. *on*)
middle [vehicles]	2 sets: plane, car (*between* vs. *middle*)[c,d]
under [crib]	2 cribs (*under* vs. *in*)
in [cabinet]	2 cabinets (*on* vs. *in*)
middle [furniture]	2 sets: cabinet, crib (*middle* vs. *between*) [c,d]
on [basket]	2 baskets (*on* vs. *in*)
between [vehicles]	2 sets: car, plane (*next to* vs. *between*) [b,d]

Note. Trials administered in the order above.

[a]Cow always placed at midpoint between two ground objects. See Fig. 3.

[b]Correct animal placed between two ground objects but closer to (next to) one. Incorrect animal also placed next to a single ground object, but not between the two ground objects. See Fig. 3.

[c]Correct animal placed at the midpoint between two ground objects. Incorrect animal placed between two ground objects, but closer to one (not the midpoint). See Fig. 3.

[d]Ground objects described as "furniture", to more easily refer to both types of ground objects.

Appendix B

Midpoint Task Encoding Strategies (Beacon, Vector, Perseverative, and Other)

Coding. Children's incorrect responses on the Midpoint Task were further coded into four error types, corresponding to particular encoding strategies, as discussed earlier: *Beacon*, *Vector*, *Perseverative*, and *Other*. In *Beacon* errors, children searched at or near one of the flags. In *Vector* errors, children searched at the same distance and direction from one of the flags as at training (i.e., six inches from one of the flags). However, children may have represented direction either as left or right (e.g., 6 in. left of the red flag) or with respect to the other flag (e.g., 6 in. from the red flag in the direction of the blue flag), or children may have made left/right errors in encoding the direction (see for example Huttenlocher et al., 1994). Therefore, we coded searches the same distance from each flag in either direction as Vector errors. In *Perseverative* errors, children searched for the treasure chest at the same location within the finding box it was found on the previous trial. Finally, errors that did not fall into one of the above categories were coded as *Other* errors. Other errors would have accounted for cases where children were genuinely using some other strategy or searching randomly. However, they may also have included cases where children were in fact using an identified encoding strategy, but they did not search within the narrow margins of our coding criteria (1.5 in. on either side of the location). Narrow margins were necessary to avoid overlap between regions, in order to discriminate different strategies from each other.

Default strategy use and strategy revision after feedback. Children's default strategy use was assessed on the first expanded trial, and potential strategy revision after feedback was assessed by looking at the last expanded trial (Table 6).

Other (i.e., uncategorizable) errors were very common. Of the incorrect, but still systematic, strategies children could have employed, children of all ages were most likely to use a Vector strategy on both the first and last expanded trial. Beacon and Perseverative errors were very rare. The overall pattern of strategy use on the first and last expanded trials was fairly consistent, suggesting that children did not dramatically change their strategies – at least with respect to these categories – across the task.

Overall performance. To assess children's overall performance, we analyzed children's proportion correct out of all seven test trials (some children did not complete all 7 trials). Among incorrect-but-systematic encoding strategies, Vector responses were the most common, mirroring the pattern seen on the first and last expanded trials. However, children's most common incorrect responses were Other errors (Table 7).

To examine how these response patterns changed over development, the mean proportion of each type of response was correlated with age in months. Older children were significantly less likely to make Vector, $r = -.50$, $p < .001$, and Other responses, $r = -.58$, $p < .001$, but age was not related to the less frequent Beacon, $r = -.14$, $p = .34$, or Perseverative responses, $r = .02$, $p = .88$.

Table 6
Encoding strategies on first (Test Trial 2) and last (Test Trial 8) expanded trial.

	Trial	3-year-olds	4-year-olds	5-year-olds	Total
Correct	First	4 (27%)	9 (50%)	15 (88%)	28 (56%)
	Last	1 (9%)	10 (59%)	12 (75%)	23 (52%)
Beacon errors	First	0 (0%)	0 (0%)	0 (0%)	0 (0%)
	Last	1 (9%)	0 (0%)	0 (0%)	1 (2%)
Vector errors	First	5 (33%)	3 (17%)	1 (6%)	9 (18%)
	Last	2 (18%)	2 (12%)	1 (6%)	5 (11%)
Perseverative errors	First	1 (7%)	0 (0%)	0 (0%)	1 (2%)
	Last	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Other errors	First	5 (33%)	6 (33%)	1 (6%)	12 (24%)
	Last	7 (64%)	5 (29%)	3 (19%)	15 (34%)
n	First	15	18	17	50
	Last	11	17	16	44

Table 7
Mean Proportion of Responses on Midpoint Task by Strategy Type.

	3-year-olds *M* (*SD*)	4-year-olds *M* (*SD*)	5-year-olds *M* (*SD*)	Total *M* (*SD*)
Out of Total Responses[Ω]	N = 15	N = 18	N = 17	N = 50
Correct	0.37 (0.24)[b,c]	0.65 (0.30)[a]	0.83 (0.16)[a]	0.63 (0.30)
Beacon errors	0.02 (0.07)	0.00 (0.00)	0.00 (0.00)	0.01 (0.04)
Vector errors	0.21 (0.17)[c]	0.10 (0.18)	0.03 (0.08)[a]	0.11 (0.17)
Perseverative errors	0.07 (0.10)	0.05 (0.08)	0.06 (0.07)	0.06 (0.08)
Other errors	0.33 (0.18)[c]	0.19 (0.18)	0.08 (0.14)[a]	0.20 (0.19)
Out of Total Errors	N = 15	N = 14	N = 12	N = 41
Beacon errors	0.02 (0.07)	0.00 (0.00)	0.00 (0.00)	0.01 (0.04)
Vector errors	0.33 (0.24)	0.20 (0.27)	0.22 (0.41)	0.25 (0.31)
Perseverative errors	0.12 (0.17)	0.24 (0.36)	0.44 (0.45)	0.26 (0.36)
Other errors	0.52 (0.25)	0.56 (0.34)	0.33 (0.40)	0.48 (0.33)

Note: These means are for all seven test trials, including the first expanded trial.
[Ω] Bonferroni corrected pairwise comparisons, $p < .05$.
[a] Different from 3-year-olds.
[b] Different from 4-year-olds.
[c] Different from 5-year-olds.

Appendix C

Language Task – Middle and Between

Our data do not allow us to distinguish understanding of *middle* and *between* with confidence and should be interpreted cautiously. Prior work on the acquisition of *between* suggests that it is acquired later than many other spatial prepositions, with both comprehension and production showing a protracted trajectory (Durkin, 1981; Hund et al., 2017; Johnston & Slobin, 1979; Weber et al., 2018; Weist, Lyytinen, Wysocka, & Atanassova, 1997). Very little work has examined the acquisition of *middle* or compared the acquisition of *middle* and *between* (but see Foster & Hund, 2012; Hund et al., 2017) (Table 8).

Overall, children's comprehension and production of both terms improved from 3- to 5-years-old. Children at all ages were more likely to produce *middle* than *between* on the production trials. This was an appropriate response, because the production trials showed a figure at the midpoint between two ground objects. However, it could also reflect the fact that children are generally more likely to produce *middle* than *between*.

Children's comprehension, measured by both the placement comprehension trials and forced-choice comprehension trials, showed a steady increase in understanding of *between* and *middle* from 3- to 5-years-old. Performance on the placement comprehension trials was better for *between* than for *middle* at all ages, though we note that the greater precision needed to accurately place the figure on *middle* trials than *between* trials means that children would have been more likely to get *between* trials correct by chance. On the forced-choice comprehension trials, only the 5-year-olds showed better understanding of *between* than *middle*.

Table 8
Mean Proportion Correct on *Middle* and *Between* Trials.

Subtest	3-year-olds		4-year-olds		5-year-olds		Total	
	M (*SD*)		*M* (*SD*)		*M* (*SD*)		*M* (*SD*)	
	middle	*between*	*middle*	*between*	*middle*	*between*	*middle*	*between*
Production[a]								
# children	3	1	10	2	11	4	24	7
Production	0.25 (0.45) n = 12		0.58 (0.49) n = 18		0.85 (0.34) n = 17		0.60 0.48) n = 47	
Comprehension								
Placement	0.38 (0.42) n = 13	0.68 (0.46) n = 14	0.79 (0.36) n = 17	0.85 (0.34) n = 17	0.79 (0.36) n = 17	1.00 (0.00) n = 17	0.68 (0.41) n = 47	0.85 (0.34) n = 48
Forced Choice	0.22 (0.26) n = 9	0.20 (0.26) n = 10	0.50 (0.48) n = 16	0.53 (0.43) n = 16	0.44 (0.48) n = 16	0.78 (0.41) n = 16	0.41 (0.43) n = 41	0.55 (0.44) n = 42

[a] *Middle* and *between* were not tested separately in the production section; the figure was always placed at the midpoint between two objects, so both *between* and *middle* were considered correct responses for these trials. The first row lists the number of children who produced each term on at least one production trial.

References

Ankowski, A. A., Thom, E. E., Sandhofer, C. M., & Blaisdell, A. P. (2012). Spatial language and children's spatial landmark use. *Child Development Research,* 1–14.

Balcomb, F., Newcombe, N. S., & Ferrara, K. (2011). Finding where and saying where: Developmental relationships between place learning and language in the first year. *Journal of Cognition and Development, 12*(3), 315–331.

Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology, 35*, 99–134.

Bushnell, E. W., McKenzie, B. E., Lawrence, D. A., & Connell, S. (1995). The spatial coding strategies of one-year-old infants in a locomotor search task. *Child Development, 66*(4), 937–958.

Carey, S. (2009). *The origin of concepts.* Oxford University Press.

Casasola, M. (2005). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology, 41*(1), 183–192.

Casasola, M., & Bhagwat, J. (2007). Do novel words facilitate 18-month-olds' spatial categorization? *Child Development, 78*(6), 1818–1829.

Casasola, M., Bhagwat, J., & Burke, A. S. (2009). Learning to form a spatial category of tight-fit relations: How experience with a label can give a boost. *Developmental Psychology, 45*(3), 711–723.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.). *Cognitive development and the acquisition of language* (pp. 65–110). New York: Academic Press.

Collett, T. S., Cartwright, B. A., & Smith, B. A. (1986). Landmark learning and visuo-spatial memories in gerbils. *Journal of Comparative Physiology A, 158*(6), 835–851.

DeLoache, J. S., & Brown, A. L. (1983). Very young children's memory for the location of objects in a large-scale environment. *Child Development, 54*(4), 888–897.

Durkin, K. (1981). Aspects of late language acquisition: School children's use and comprehension of prepositions. *First Language, 2*(4), 47–59.

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

Foster, E. K., & Hund, A. M. (2012). The impact of scaffolding and overhearing on young children's use of the spatial terms *between* and *middle*. *Journal of Child Language, 39*(2), 338–364.

Frick, A., & Newcombe, N. S. (2012). Getting the big picture: Development of spatial scaling abilities. *Cognitive Development, 27*(3), 270–282.

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development, 49*, 988–998.

Gentner, D. (2003). *Why we're so smart. In Language in mind. Advances in the study of language and thought.* Cambridge, MA: MIT Press195–235.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775.

Gentner, D. (2016). Language as cognitive toolkit: How language supports relational thought. *The American Psychologist, 71*(8), 650–657.

Gentner, D., & Christie, S. (2012). Language and cognition in development. In M. M. Spivey, K. McRae, & M. Joanisse (Eds.). *The Cambridge handbook of psycholinguistics* (pp. 653–673). Cambridge, England: Cambridge University Press.

Gentner, D., & Loewenstein, J. (2002). Relational language and relational thought. In E. Amsel, & J. P. Byrnes (Eds.). *Language, literacy, and cognitive development: The development and consequences of symbolic communication* (pp. 87–120). Mahwah, NJ: Lawrence Erlbaum Associates.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*(6), 297–301.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman, & J. P. Byrnes (Eds.). *Perspectives on language and thought* (pp. 225–277). Cambridge: Cambridge University Press.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*(3), 277–300.

Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. (2009). Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science, 33*, 1343–1382.

Gentner, D., Özyürek, A., Gürcanli, Ö., & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition, 127*(3), 318–330.

Goldin-Meadow, S. (1993). When does gesture become language? A study of gesture used as a primary communication system by deaf children of hearing parents. In K. R. Gibson, & T. Ingold (Eds.). *Tools, language and cognition in human evolution* (pp. 63–85). Cambridge, UK: Cambridge University Press.

Haun, D. B. M., Rapold, C. J., Call, J., Janzen, G., & Levinson, S. C. (2006). Cognitive cladistics and cultural override in Hominid spatial cognition. *Proceedings of the National Academy of Sciences, 103*(46), 17568–17573.

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition, 79*(3), 263–299.

Hund, A. M., Bianchi, L. J., Winner, J. F., & Hesson-McInnis, M. S. (2017). Complex spatial language improves from 3 to 5 years: The role of prompting and overhearing in facilitating direction giving using *between* and *Middle*. *Cognitive Development, 43*, 170–181.

Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology, 27*(2), 115–147.

Jacques, S., & Zelazo, P. D. (2005). 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* (pp. 144–162). Oxford, UK: Oxford University Press.

Jamrozik, A., & Gentner, D. (2013). Relational labels can improve relational retrieval. M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.). *Proceedings of the 35th Annual Conference of the Cognitive Science Society,* 651–656.

Jamrozik, A., & Gentner, D. (2015). Well-hidden regularities: Abstract uses of *in* and *on* retain an aspect of their spatial meaning. *Cognitive Science, 39*, 1881–1911.

Jeong, Y., Levine, S. C., & Huttenlocher, J. (2007). The development of proportional reasoning: Effect of continuous versus discrete quantities. *Journal of Cognition and Development, 8*(2), 237–256.

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.

Johnston, J. R., & Slobin, D. I. (1979). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language, 6*(3), 529–545.

Jones, J. E., Antoniadis, E., Shettleworth, S. J., & Kamil, A. C. (2002). A comparative study of geometric rule learning by nutcrackers (Nucifraga columbiana), pigeons (Columba livia), and jackdaws (Corvus monedula). *Journal of Comparative Psychology, 116*(4), 350–356.

Kamil, A. C., & Jones, J. E. (1997). The seed-storing corvid Clark's nutcracker learns geometric relationships among landmarks. *Nature, 390*(6657), 276–279.

Kamil, A. C., & Jones, J. E. (2000). Geometric rule learning by Clark's nutcrackers (Nucifraga columbiana). *Journal of Experimental Psychology Animal Behavior Processes, 26*(4), 439–453.

Kersten, A. W., & Earles, J. L. (2004). Semantic context influences memory for verbs more than memory for nouns. *Memory & Cognition, 32*(2), 198–211.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension switching task. *Developmental Science, 6*, 449–476.

Learmonth, A. E., Nadel, L., & Newcombe, N. S. (2002). Children's use of landmarks: Implications for modularity theory. *Psychological Science, 13*(4), 337–341.

Levelt, W. J. M. (2005). Habitual perspective. B. G. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.). *Proceedings of the 27th Annual Conference of the Cognitive Science Society*.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50*(4), 315–353.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*(12), 1077–1083.

MacDonald, S. E., Spetch, M. L., Kelly, D. M., & Cheng, K. (2004). Strategies in landmark use by children, adults, and marmoset monkeys. *Learning and Motivation, 35*(4), 322–347.

Marsh, H. L., Spetch, M. L., & MacDonald, S. E. (2011). Strategies in landmark use by orangutans and human children. *Animal Cognition, 14*, 487–502.

Miller, S. E., & Marcovitch, S. (2011). Toddlers benefit from labeling on an executive function search task. *Journal of Experimental Child Psychology, 108*, 580–592.

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

Newcombe, N., & Huttenlocher, J. (2003). *Making space: The development of spatial representation and reasoning.* MIT Press.
Newcombe, N., Huttenlocher, J., Drummey, A. B., & Wiley, J. G. (1998). The development of spatial location coding: Place learnng and dead reckoning in the second and third years. *Cognitive Development, 13*(2), 185–200.
Newcombe, N. S., Uttal, D. H., & Sauter, M. (2013). Spatial development. In P. D. Zelazo (Ed.). *Oxford handbook of developmental psychology* (pp. 564–590). New York: Oxford University Press.
Potì, P., Kanngiesser, P., Saporiti, M., Amiconi, A., Bläsing, B., & Call, J. (2010). Searching in the middle—Capuchins' (Cebus apella) and bonobos' (Pan paniscus) behavior during a spatial search task. *Journal of Experimental Psychology: Animal Behavior Processes, 36*(1), 92–109.
Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Developmental Science, 14*(6), 1417–1430.
Pyers, J. E., Shusterman, A., Senghas, A., Spelke, E. S., & Emmorey, K. (2010). Evidence from an emerging sign language reveals that language supports spatial cognition. *Proceedings of the National Academy of Sciences, 107*(27), 12116–12120.
Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. *Cognitive Development, 18*(4), 511–531.
Shusterman, A., Lee, S. A., & Spelke, E. S. (2011). Cognitive effects of language on human navigation. *Cognition, 120*(2), 186–201.
Son, J. Y., Smith, L. B., Goldstone, R. L., & Leslie, M. (2012). The importance of being interpreted: Grounded words and children's relational reasoning. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00045.
Sovrano, V. A., Bisazza, A., & Vallortigara, G. (2007). How fish do geometry in large and in small spaces. *Animal Cognition, 10*(1), 47–54.
Spencer, J. P., & Hund, A. M. (2002). Prototypes and particulars: Geometric and experience-dependent spatial categories. *Journal of Experimental Psychology General, 131*(1), 16–37.
Spencer, J. P., & Hund, A. M. (2003). Developmental continuity in the processes that underlie spatial recall. *Cognitive Psychology, 47*(4), 432–480.
Spetch, M. L., & Parent, M. B. (2006). Age and sex differences in children's spatial search strategies. *Psychonomic Bulletin & Review, 13*(5), 807–812.
Spetch, M. L., Cheng, K., & MacDonald, S. E. (1996). Learning the configuration of a landmark array: I. Touch-screen studies with pigeons and humans. *Journal of Comparative Psychology, 110*(1), 55–68.
Spetch, M. L., Cheng, K., MacDonald, S. E., Linkenhoker, B. A., Kelly, D. M., & Doerkson, S. R. (1997). Use of landmark configuration in pigeons and humans: II. Generality across search tasks. *Journal of Comparative Psychology, 111*(1), 14–24.
Spinillo, A. G., & Bryant, P. (1991). Children's proportional judgments: The importance of "half". *Child Development, 62*(3), 427–440.
Sturz, B. R., & Katz, J. S. (2009). Learning of absolute and relative distance and direction from discrete visual landmarks by pigeons (Columba livia). *Journal of Comparative Psychology, 123*(1), 90–113.
Sutton, J. E., Olthof, A., & Roberts, W. A. (2000). Landmark use by squirrel monkeys (Saimiri sciureus). *Animal Learning & Behavior, 28*(1), 28–42.
Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology, 78*, 57–77.
Tomasello, M. (1987). Learning to use prepositions: A case study. *Journal of Child Language, 14*(1), 79–98.
Uttal, D. H., Gregg, V. H., Tan, L. S., Chamberlin, M. H., & Sines, A. (2001). Connecting the dots: Children's use of a systematic figure to facilitate mapping and search. *Developmental Psychology, 37*(3), 338–350.
Uttal, D. H., Sandstrom, L. B., & Newcombe, N. S. (2006). One hidden object, two spatial codes: Young children's use of relational and vector coding. *Journal of Cognition and Development, 7*(4), 503–525.
Wagner, K., Jergens, J., & Barner, D. (2018). Partial color word comprehension precedes production. *Language Learning and Development, 14*, 241–261.
Weber, J. M., Miller, H. E., & Ou, L. (2018). Children's representations of five spatial terms. T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.). *Proceedings of the 40th Annual Conference of the Cognitive Science Society,* 1648–1653.
Weist, R. M., Lyytinen, P., Wysocka, J., & Atanassova, M. (1997). The interaction of language and thought in children's language acquisition: A crosslinguistic study. *Journal of Child Language, 24*(1), 81–121.