An adaptive cue combination model of spatial reorientation

Yang Xu (yang_xu_ch@berkeley.edu)   Terry Regier (terry.regier@berkeley.edu)
Department of Linguistics
University of California, Berkeley

Nora Newcombe (newcombe@temple.edu)
Department of Psychology
Temple University

Abstract

Humans use a variety of strategies to reorient in space. There are diverging views on whether spatial reorientation relies on an encapsulated geometric module, an associative mechanism or an adaptive combination of different cues. We test these proposals with a computational model that predicts human behavior in reorientation. By analyzing existing data from multiple sources, we show evidence for an adaptive view of reorientation that combines information from geometry, direction and language. Our work opens up opportunities to understand the interactive strategies of human reorientation.

Keywords: spatial reorientation; geometric module; room-size effect; associative model; adaptive cue combination; probabilistic models.

Introduction

The spatial world provides many cues to where things are. Because these cues usually converge, one way of thinking about how they support navigation is Bayesian combination, using weights based on saliency, reliability and validity. When there are large discrepancies between cues, perhaps resulting from encoding errors or forgetting, choices are made between cues according to the same considerations (Cheng, Huttenlocher, & Rieser, 2007; Huttenlocher, Hedges, & Duncan, 1991).

An alternative perspective is that spatial cues are modularly separable. Research on behavior when organisms are disoriented and internally-generated spatial cues are not useful initially seemed to support modularity, because geometrical cues were used while potentially useful featural cues were not (Cheng, 1986; Gallistel, 1990; Hermer & Spelke, 1994, 1996). While human adults do use featural cues, Hermer-Vazquez, Spelke and Katsnelson (1999) argued that young children and non-human species share a geometric module for reorientation, later punctured by spatial language.

The modularity hypothesis has attracted much attention, but it has become clear that it cannot account for many aspects of the expanding data set. One prominent problem is the room-size effect. Geometry is more likely to be used in small spaces and features are more likely to be used in large spaces, for children (Learmonth, Newcombe, & Huttenlocher, 2001; Learmonth, Nadel, & Newcombe, 2002), adults (Ratliff & Newcombe, 2008b), fish (Sovrano, Bisazza, & Vallortigara, 2007), chicks (Chiandetti, Regolin, Sovrano, & Vallortigara, 2007; Sovrano & Vallortigara, 2006; Vallortigara, Feruglio, & Sovrano, 2005), and pigeons (Kelly, Spetch, & Heth, 1998). In addition, short-term experience

with the usefulness of a feature cue changes the behavior of young children (Twyman, Friedman, & Spetch, 2007), human adults (Ratliff & Newcombe, 2008a) and pigeons (Kelly & Spetch, 2004). Further, rearing environment changes weighting of geometry and features, at least for convict fish (Brown, Spetch, & Hurd, 2007) and mice (Twyman, Newcombe, & Gould, 2013), although not chicks (Chiandetti & Vallortigara, 2008, 2010). Cheng (2008) suggested abandoning a modularity approach.

Non-modular approaches to the reorientation data have been proposed; for an overview, see Cheng, Huttenlocher, and Newcombe (2013). One computational approach is an associative model (Miller & Shettleworth, 2007; Miller, 2009). A Bayesian alternative called the adaptive combination model has been proposed (e.g. Newcombe & Huttenlocher, 2006) but not yet computationally specified. The purpose of this paper is to specify such a computational model, test it against and cross-predict existing data, and compare it to alternatives.

Computational model

We assume that an agent (e.g. a person or an animal) is inside a closed space (e.g. a room), has seen an object being hidden in one of a finite number of possible locations within that space (e.g. one of the corners of a room), and is then disoriented within that space. We model the agent's belief concerning the location of the target object t as jointly influenced by the three cues G, D, and L, as illustrated in Figure 1.
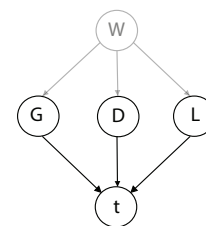

Figure 1: Graphical representation of the model. G: geometry; D: direction; L: language; t: choice of target location; W: world (not explicitly modeled).

G (geometry) is the geometric shape of the space, e.g. the lengths of the walls, and the angles they form where they meet. D (direction) assumes that the agent has oriented, or polarized, toward a distinctive feature-bearing landmark, such as a colored wall, and encodes the angular distance of the

target object from the agent's front when the agent is facing that landmark. Note that purely associative reorientation—commonly understood as locating a target based on its spatial coincidence with a landmark—is a special case of D in which the angular distance between landmark and target is zero. The directional effect could also be described allocentrically rather than egocentrically. L (language) is the ability to use spatial language to assist in locating the target. These cues are themselves derived in some way from the structure of the world (W), but our current work does not explicitly model how world structure gives rise to these cues, and for this reason that portion of the model is grayed out in the figure. We define C as the set of available cues $C = \{G, D, L\}$ and capture reorientation and selection of the target location as probabilistic inference over possible target locations given this combined set of cues. Using Bayes' rule, we express the posterior belief concerning target location t as:

$$p(t|C) \propto p(C|t)p(t) = p(G,D,L|t)p(t) \qquad (1)$$

Assuming initial ignorance about possible locations, i.e. a uniform prior $p(t)$, the posterior is proportional to the likelihood $p(G,D,L|t)$. Since each cue contributes independently to reorientation, we further decompose the likelihood by separating out the cues using this independence assumption:

$$
\begin{aligned}
p(G,D,L|t) &= p(G|t)p(D|t)p(L|t) \qquad (2)\\
&\propto f(t|G)p(G)f(t|D)p(D)f(t|L)p(L)
\end{aligned}
$$

Thus, the overall likelihood $p(G,D,L|t)$ depends on how likely it is that the target t will be predicted by each cue $f(t|\cdot)$, weighted by prior belief in each cue $p(\cdot)$. For all of our present analyses, we assume priors on cues are uniform, although our framework makes it possible to encode prior preferences over cues. Thus what remains is to specify the cue likelihood term $f(t|\cdot)$ for each cue.

For each cue, we represent its likelihood by a function that expresses how likely the target is at each given location. This function encodes the information that each cue provides about target location, and includes a signal component and a noise component. We specify a single noise parameter δ for all cues, and this noise parameter diminishes with experience or training. We use age as a proxy for experience. Thus for any given cue with no *a priori* training, this model holds that children are more likely than adults to receive noisy information from that cue, which is one reason children are more likely to make incorrect choices concerning target location.

For the G cue, the likelihood function specifies a signal component of 1 if that cue predicts a certain location i as the target. For example, when reorienting in a rectangularly shaped room to a target at one of the four corners, G has a value of 1 on the target corner and on the corner diagonally opposite it, because those are the two locations that are identical to the target location when considering only geometric information. The likelihood also specifies a noise compo-

nent that scales inversely with age subject to a multiplicative weighting parameter w:

$$f(t=i|G) = \begin{cases} 1 & \text{(if } G \text{ predicts } i)\\ \delta = \frac{1}{w \cdot age} & \text{(otherwise)} \end{cases} \qquad (3)$$

The likelihood of the D cue has an identical noise component but a different signal component. We assume that D has a signal component that points only to the target object, but the strength of this cue depends on the salience of the landmark (e.g. the salience of a colored wall). This signal component would be perfectly informative in a high-salience condition, but uninformative in a low-salience condition. This dependence on salience is motivated by previous findings that a small enclosure can degrade reorientation performance despite the presence of a landmark feature, especially in very young children (Hermer & Spelke, 1996). We assume that small room size reduces the salience of the landmark feature. We formalize this interaction between salience and the D cue using a probabilistic OR:

$$f(t=i|D) = \begin{cases} 1-(1-s)(1-\delta) & \text{(if } D \text{ predicts } i)\\ \delta & \text{(otherwise)} \end{cases} \qquad (4)$$

Here s represents salience and can range from 0 (no salience) to 1 (full salience). When s is 0, this likelihood function is entirely flat and uninformative. Here, the salience parameter s is assumed to be 0 in a small room (area no greater than 24 square feet, motivated by previous work by Hermer & Spelke, 1996), and 1 otherwise. Although we have made this binary assumption for simplicity, future work can explore independent empirical measures of landmark salience for rooms of different sizes in a systematic way.

The likelihood of the L cue takes a similar form, but depends on the availability of spatial language rather than salience:

$$f(t=i|L) = \begin{cases} 1-(1-a(L))(1-\delta) & \text{(if } L \text{ predicts } i)\\ \delta & \text{(otherwise)} \end{cases} \qquad (5)$$

This cue points to the target and becomes available when an agent is fluent and has access to spatial language, which is determined by the function $a(L)$. This function ranges from 0 (no access to spatial language) to 1 (full access). We assume a value of 0 (no access) for participants under the age of 6, a value of 1 (full access) for those who are 6 years of age or older and are not experimentally prevented from using language, and a range of values between 0.5 and 1 for participants 6 years of age or older who are being experimentally prevented from using language, for example through a concurrent verbal interference task. These assumptions are motivated by previous work suggesting that spatial language helps children of age 6 but not younger in reorientation (Hermer-Vazquez, 1997), and that verbal shadowing degrades reorientation performance in adults (Hermer-Vazquez, Spelke, & Katsnelson, 1999).

Table 1: Summary of data sets. LNN02: Learmonth, Nadel & Newcombe (2002); HS96: Hermer & Spelke (1996); RN08: Ratliff & Newcombe (2008); NRST10: Newcombe, Ratliff, Shallcross & Twyman (2010); TFS07: Twyman, Friedman & Spetch (2007).

	Theoretical implication	Phenomenon	Source
1	Geometrical module is not impenetrable	≤6-year-olds use landmark to reorient	LNN02
2	Salience (enclosure size) matters in reorientation	<6-year-olds fail to use landmark in small yet not large enclosure	LNN02,HS96
3	Spatial language helps reorientation	Adults show degraded performance under verbal shadowing	RN08,HS96
4	Reorientation is more than associative	<6-year-olds use distant landmark to reorient	NRST10

Data

To assess our model, we used empirical data from several existing published sources. We chose these data sets to account for the breadth of phenomena reported in the reorientation literature and maintain analytical consistency across studies. For example, we restricted our analysis to experimental conditions where the landmark is a wall. Table 1 summarizes the empirical data relevant to our analyses and their theoretical implications.

Model evaluation and results

We proceed in two steps. First, we fit our model using data from existing studies and use it to predict studies from which it was not derived. We show that our GDL model outperforms an existing associative model (Miller, 2009; Miller & Shettleworth, 2007) in accounting for associative cue combination, i.e. when the target is spatially coincident with a feature-bearing landmark. Next, we show that our model also predicts non-associative cue combination, i.e. when the target is at some distance from the landmark.

Associative cue combination

We first assess whether the GDL model accounts for associative cue combination. We begin by drawing on a representative data set that spans across age groups (Learmonth et al., 2002). This set was used by Miller (2009), so we can directly compare our model to Miller's associative model.

The experimental procedures follow Learmonth et al. (2002); we explain these briefly. All experiments were carried out in an enclosed rectangular room. The size of the room differed across two conditions: large (8-×12-ft) and small (4-×6-ft). In both cases, one of the shorter walls had a blue curtain that served as a landmark (the "blue-wall condition"). Figure 2a illustrates this experimental setup. The target was placed at a corner that coincides with the blue wall for a given trial. The four corners of the room were referred to as the C (correct), R (rotationally identical or invariant), N (near) and F (far) corners. Successful reorientation depends on associating the target with one corner of the blue wall.[1] Three groups

of children from 3 to 6 years of age participated in the reorientation task. In each trial, participants were shown the target location, disoriented and then asked to point to the target. Figure 2b-c shows the average empirical choice probabilities along with model fits.

We modeled these data using our cue combination (GDL) model as follows. The G cue was encoded by a likelihood function that has signals at the C and R corners and noise at the remaining two corners. This captures the fact that the C and R corners are geometrically identical or rotationally invariant (Hermer & Spelke, 1996); thus geometry ambiguously predicts target location. The D cue has a signal at C and noise elsewhere. This captures the fact that the target coincides with the landmark blue wall. However, the signal component for D is dependent on salience, which in this case is assumed to be determined by room size. In the small room, salience $s = 0$; in the large room, salience $s = 1$. For the L cue, we assumed full access to spatial language ($L(a) = 1$) for children 6 years of age and no access ($L(a) = 0$) otherwise. The only free parameter in the model is w, which determines how quickly noise reduces as a function of age; we fit this parameter to the data through grid search between 0.01 and 1.0 in steps of 0.01.

We compared the choice probabilities estimated from our GDL model to those from Miller's associative model (2009) [2] and a baseline model that involves only the geometrical cue with no cue combination. Figure 2b-c summarizes the results. In the large-room condition (Figure 2b), performance is almost indistinguishable between the GDL ($r^2 = 0.98$) and Miller ($r^2 = 0.97$) models, and both outperformed the baseline model ($r^2 = 0.40$). This result is expected because in the large room, the language cue L (which is absent from Miller's model) is redundant with the directional cue D (which is present in Miller's model as an associative cue), and both cues correctly point to the target; thus both of these models have this critical information, and the geometric baseline model lacks it. In the small-room condition (Figure 2c), both Miller's ($r^2 = 0.57$) and baseline ($r^2 = 0.40$) models perform poorly in explaining the data, but our proposed model accounts for the choice pattern substantially better ($r^2 = 0.96$). In particular, Miller's model predicts higher choice probability on the correct corner than other corners for all age groups, hence it fails to predict the qualitative difference in

---

[1]Learmonth et al. (2002) also included a condition in which the target was distant from the landmark, i.e. non-associative, but did not report any significant difference in behavior between associative and non-associative conditions. Therefore we focus on the associative condition here, although our result holds equally for the non-associative case.

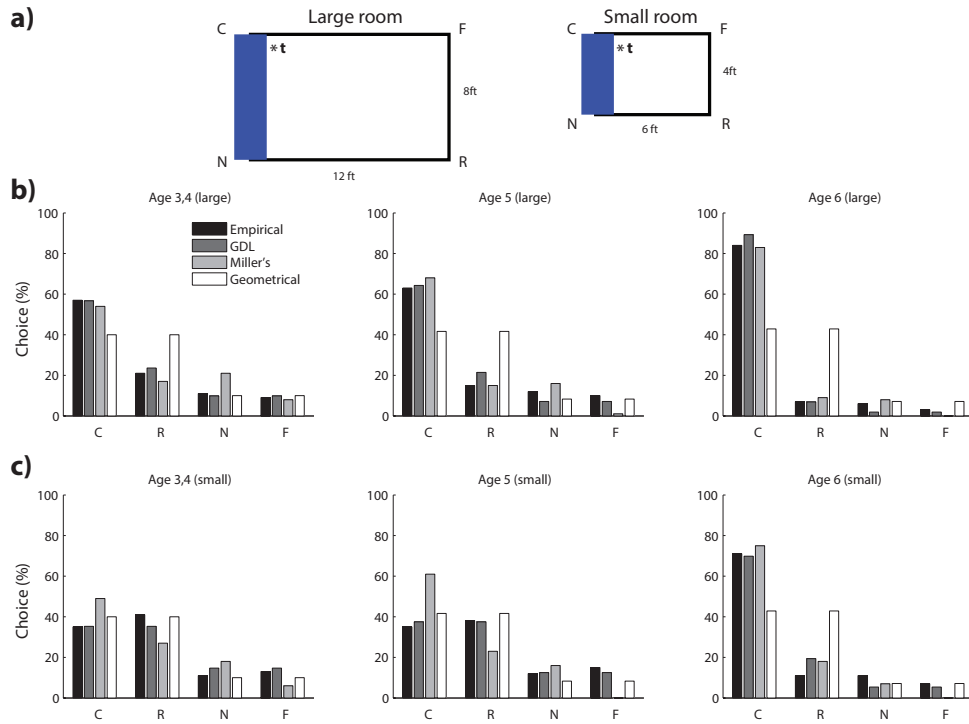[2]Fits for the associative model are extracted from Miller (2009).

Figure 2: Summary of analyses on associative cue combination in children. a) Large and small rectangular rooms (not-to-scale) used in the experiments by Learmonth et al. (2002). "*t" marks target location. b) Empirical and model-estimated choice probabilities from children for the large room. c) Similar choice probabilities for the small room.

performance between age groups < 6 and $= 6$. On the other hand, the geometry-only model fails to capture the superior performance in age group 6 altogether. This is because neither model captures the presumed influence of language when salience is weak in the small room. Our model, however, does explicitly capture this interaction between language and salience, and best explains these empirical findings that contrast reorientation performance across ages between large and small enclosures.

To further assess our model, we used the parameters estimated from the previous experiment to predict data from a separate study conducted with adults (Ratliff & Newcombe, 2008a).[3] The specific paradigm we tested is consistent with the blue-wall experiment described earlier: adults were disoriented in identically sized large and small rooms. However in this study, one of the conditions involved verbal shadowing, intended to render language partially inaccessible during reorientation (see also Hermer-Vazquez et al., 1999). Figure 3 (first column) presents data from Ratliff & Newcombe (2008a). It can be seen that verbal shadowing appears to somewhat reduce the rate of correct responding, and that this reduction is (significantly) smaller in the large room than in the small room. We interpret the greater resilience of correct responding under verbal shadowing in the large room as resulting from the greater salience of the landmark in the large

---

[3]We took adults' age to be 18 and used the parameter value for *w* from the previous analysis to determine the noise term.

room (supporting the *D* cue), despite the experimentally-induced degradation of the *L* cue in both conditions. Figure 3 (columns 2-4) presents the response patterns predicted by the same three models described above (we varied the availability of language $a(L)$ in the *L* cue over the range 0.5 to 1 (in steps of 0.1) to capture different degrees of language inaccessibility; the bar graphs for the GDL model show the average prediction within that range). Our model again outperforms ($r^2 = 0.94$) the competing models by capturing the interaction between salience (here, room size) and language, whereas the competing models do not account for this interaction.

Taken together, these two sets of results suggest our cue combination model accounts well for association-based cue combination in reorientation.

Non-associative cue combination

We next assess whether our model also accounts for cue combination beyond association—a phenomenon that is not captured by previous computational models. In particular, we focus on directional cue combination as reported by Newcombe et al. (2010) .

In this set of experiments, the reorientation task was conducted in an octagonal room. Similar to the previous experiments, one of the walls had a red curtain that served as a landmark. However in contrast with the previous experiments, the target location did not coincide with any of the corners covered by the landmark, but instead was one of the distant
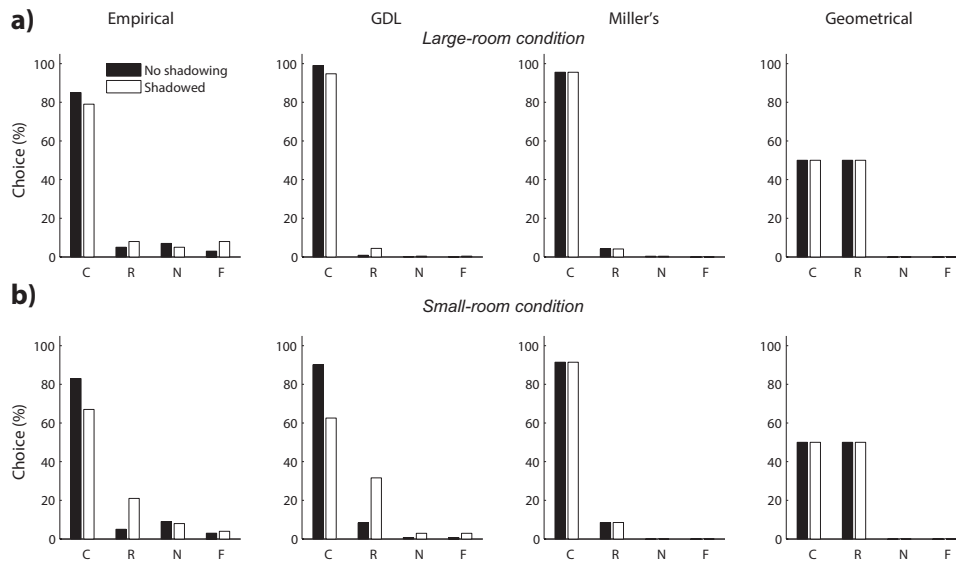
Figure 3: Summary of analyses on associative cue combination in adults. a) Empirical and model-estimated choice probabilities under control and verbal-shadowing conditions for the large room. b) Similar choice probabilities for the small room.

corners as illustrated in Figure 4a (leftmost panel). In other words, successful reorientation in this case would depend on mapping from a distant landmark to a target, or direction. The room was 75 square feet in area, which is close to the size of the previous large rectangular room. Two of the corners (GW) were geometrically identical to the correct corner (C). One corner (GF) was both geometrically identical and coincides with the featured landmark. The remaining three corners were error corners (EW). Two groups of children (ages 3 and 5) participated in the experiment. Figure 4b (right 2 panels) shows the empirical choice probabilities, and model estimates of them.

We used the GDL model to account for these empirical data. Specifically, the G cue had signal components at all geometrically invariant corners (C,GF,GW and GW) and noise δ elsewhere. The directional D cue had a signal component at C and noise elsewhere. The L cue is uninformative in this case because both age groups for the octagonal experiments are below 6, which renders a flat likelihood. To allow for a fully predictive (parameter-free) assessment, we used the noise values for 3- and 5-year-olds from the blue-wall experiments, which are independent of the current analysis.

We compared our GDL model against two baseline models that incorporate either the directional cue alone, or the geometrical cue alone, with no cue combination. Figure 4b summarizes the predicted choice probabilities from all three models. These results show that our GDL model ($r^2 = 0.87$) accounts for empirical data across two age groups better than the direction- ($r^2 = 0.68$) or geometry- ($r^2 = 0.32$) based model. The directional model cannot distinguish between geometrically identical corners and error corners, so it over-predicts the overall erroneous choices. The geometrical model cannot distinguish between the correct corner and ge-

ometrically confounded corners, so it under-predicts choices on the target. In sum, this set of results suggests that children combine non-associative, directional information with geometry in spatial reorientation, which extends beyond an associative account and provides firmer evidence for a cue-combination view.

Conclusion

We have presented a probabilistic model of spatial reorientation based on cue combination. Our model accounts for human behavior better than existing models across different enclosure sizes and shapes, different age groups, and different landmark configurations. Despite its simplicity, our model suggests that reorientation is neither performed in an isolated module nor purely association-based, hence moving beyond that debate and helping to illuminate the richness of human spatial strategies.

Acknowledgments

References

Brown, A. A., Spetch, M. L., & Hurd, P. L. (2007). Growing in circles rearing environment alters spatial navigation in fish. *Psychological Science*, *18*(7), 569–573.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, *23*(2), 149–178.

Cheng, K. (2008). Whither geometry? Troubles of the geometric module. *TICS*, *12*(9), 355–361.

Cheng, K., Huttenlocher, J., & Newcombe, N. (2013). 25 years of research on the use of geometry in spatial reori-
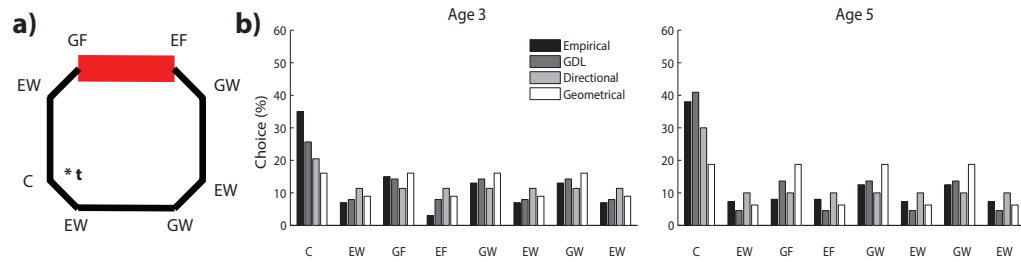
Figure 4: Summary of analyses on directional cue combination. a) Octagonal room used in the experiments by Newcombe et al. (2009). b) Empirical and model-estimated choice probabilities for 3- and 5-year-olds.

entation: A current theoretical perspective. *Psychonomic bulletin & review*, *20*(6), 1033–1054.

Cheng, K., Huttenlocher, S. J. S. J., & Rieser, J. J. (2007). Bayesian integration of spatial information. *Psychological bulletin*, *133*(4), 625.

Chiandetti, C., Regolin, L., Sovrano, V. A., & Vallortigara, G. (2007). Spatial reorientation: The effects of space size on the encoding of landmark and geometry information. *Animal cognition*, *10*(2), 159–168.

Chiandetti, C., & Vallortigara, G. (2008). Is there an innate geometric module? Effects of experience with angular geometric cues on spatial reorientation based on the shape of the environment. *Animal Cognition*, *11*(1), 139–146.

Chiandetti, C., & Vallortigara, G. (2010). Experience and geometry: Controlled-rearing studies with chicks. *Animal Cognition*, *13*(3), 463–470.

Gallistel, C. R. (1990). *The organization of learning*. The MIT Press.

Hermer, L., & Spelke, E. (1996). Modularity and development: The case of spatial reorientation. *Cognition*, *61*(3), 195–232.

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, *370*(6484), 57–59.

Hermer-Vazquez, L. (1997). Cognitive flexibility as it emerges over development and evolution: The case of two navigational tasks in humans. *Unpublished doctoral dissertation, Cornell University, Ithaca, NY*.

Hermer-Vazquez, L., Spelke, E., & Katsnelson, A. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive psychology*, *39*(1), 3–36.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*(3), 352.

Kelly, D. M., & Spetch, M. L. (2004). Reorientation in a two-dimensional environment: Do pigeons encode the featural and geometric properties of a two-dimensional schematic of a room? *J Comp Psychol*, *118*(4), 384.

Kelly, D. M., Spetch, M. L., & Heth, C. D. (1998). Pigeons' encoding of geometric and featural properties of a spatial environment. *J Comp Psychol*, *112*(3), 259.

Learmonth, A. E., Nadel, L., & Newcombe, N. (2002). Chil-

dren's use of landmarks: Implications for modularity theory. *Psychological Science*, *13*(4), 337–341.

Learmonth, A. E., Newcombe, N., & Huttenlocher, J. (2001). Toddlers' use of metric information and landmarks to reorient. *Journal of Experimental Child Psychology*, *80*(3), 225–244.

Miller, N. (2009). Modeling the effects of enclosure size on geometry learning. *Behavioural Processes*, *80*(3), 306–313.

Miller, N., & Shettleworth, S. J. (2007). Learning about environmental geometry: An associative model. *JEP: Animal Behavior Processes*, *33*(3), 191.

Newcombe, N., & Huttenlocher, J. (2006). Development of spatial cognition. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology (6th edition)* (pp. 734–776). John Wiley.

Newcombe, N., Ratliff, K., Shallcross, W., & Twyman, A. (2010). Young children's use of features to reorient is more than just associative: Further evidence against a modular view of spatial processing. *Dev Sci*, *13*, 213–220.

Ratliff, K., & Newcombe, N. (2008a). Is language necessary for human spatial reorientation? Reconsidering evidence from dual task paradigms. *Cognitive psychology*, *56*(2), 142–163.

Ratliff, K., & Newcombe, N. (2008b). Reorienting when cues conflict: Evidence for an adaptive-combination view. *Psychological Science*, *19*(12), 1301–1307.

Sovrano, V., Bisazza, A., & Vallortigara, G. (2007). How fish do geometry in large and in small spaces. *Animal cognition*, *10*(1), 47–54.

Sovrano, V., & Vallortigara, G. (2006). Dissecting the geometric module: A sense linkage for metric and landmark information in animals' spatial reorientation. *Psychological Science*, *17*(7), 616–621.

Twyman, A., Friedman, A., & Spetch, M. L. (2007). Penetrating the geometric module: Catalyzing children's use of landmarks. *Dev Psychol*, *43*(6), 1523.

Twyman, A., Newcombe, N., & Gould, T. J. (2013). Malleability in the development of spatial reorientation. *Developmental psychobiology*, *55*(3), 243–255.

Vallortigara, G., Feruglio, M., & Sovrano, V. A. (2005). Reorientation by geometric and landmark information in environments of different size. *Dev Sci*, *8*(5), 393–401.