Cognition 163 (2017) 56–66

Contents lists available at ScienceDirect

Cognition

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

Original Articles

An adaptive cue combination model of human spatial reorientation

Yang Xu [a,*], Terry Regier [a], Nora S. Newcombe [b]

[a] Department of Linguistics, Cognitive Science Program, University of California, Berkeley, CA 94720-2650, USA
[b] Department of Psychology, 318 Weiss Hall, Temple University, Philadelphia, PA 19122, USA

ARTICLE INFO

Article history:
Received 1 January 2016
Revised 17 February 2017
Accepted 28 February 2017
Available online 10 March 2017

Keywords:
Spatial reorientation
Geometric module
Room-size effect
Associative model
Adaptive cue combination
Probabilistic model

ABSTRACT

Previous research has proposed an adaptive cue combination view of the development of human spatial reorientation (Newcombe & Huttenlocher, 2006), whereby information from multiple sources is combined in a weighted fashion in localizing a target, as opposed to being modular and encapsulated (Hermer & Spelke, 1996). However, no prior work has formalized this proposal and tested it against existing empirical data. We propose a computational model of human spatial reorientation that is motivated by probabilistic approaches to optimal perceptual cue integration (e.g. Ernst & Banks, 2002) and to spatial location coding (Huttenlocher, Hedges, & Duncan, 1991). We show that this model accounts for data from a variety of human reorientation experiments, providing support for the adaptive combination view of reorientation.

Published by Elsevier B.V.

1. Introduction

The spatial world provides many cues to where things are. For example, a pirate may have buried a treasure chest five paces east from a distinctive tree and one hundred paces away from the shore. Locating the treasure often requires combining the various cues to locating the treasure in a probabilistic fashion, using appropriate weightings (Cheng, Shettleworth, Huttenlocher, & Rieser, 2007; Huttenlocher, Hedges, & Duncan, 1991). Combining cues allows for reduction of uncertainty concerning encoding and memory for individual cues, each of which might not be sufficiently informative in isolation, but which can jointly provide more precise, if sometimes biased, localization of a target. This view of perception and memory is also seen in the literature on perceptual cue integration (Berniker & Kording, 2011; Ernst & Banks, 2002; Jacobs, 2002; Knill & Pouget, 2004). It is typically formulated in terms of probabilistic inference, which provides a rational account of human behavior under uncertainty. Such a probabilistic approach has recently begun to be a focus of navigation research, especially in studies of how egocentric and allocentric systems interact with each other (e.g., Sjolund, Kelly, & McNamara, 2014; Waisman, Lucas, Griffiths, & Jacobs, 2011; Zhao & Warren, 2015). But it has not been formally specified in explaining human behavior in spatial reorientation, an area in which there have been high-profile claims of modularity and information encapsulation, to which the cue-combination view provides an important alternative. Here we address this gap by formalizing cue combination in probabilistic terms and testing it against data on the development of spatial reorientation in human children and adults.

Research on behavior when organisms are disoriented (and therefore when egocentric spatial cues are not useful) seemed initially to support modularity, because geometric cues were used while potentially useful featural cues were not, both by rats and by young children (Gallistel, 1990; Hermer & Spelke, 1994, 1996). The classic experiments were conducted in a rectangular room, in which the relative length of the walls defines two pairs of congruent corners (i.e., long wall to the left of short wall, or vice versa). Searches were directed to the correct corners as defined by geometry, but the addition of a feature such as one colored wall did not lead participants to narrow the choice to the correct corner. Although human adults do use featural cues, Hermer-Vazquez, Spelke, and Katsnelson (1999) argued that they do so only because human language allows for the combination of the output of different processing modules, a combination that they argued would not be possible without language. In this view, young children and non-human species share an ancestral 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 on human reorientation and its development. 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, 2002), adults (Ratliff & Newcombe, 2008b),

* Corresponding author at: Department of Linguistics, 1203 Dwinelle Hall, University of California, Berkeley CA 94720-2650, USA.
E-mail address: yang_xu_ch@berkeley.edu (Y. Xu).

http://dx.doi.org/10.1016/j.cognition.2017.02.016
0010-0277/Published by Elsevier B.V.

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 featural 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.

Other non-modular approaches to the development of human reorientation have been proposed besides adaptive combination; for an overview, see Cheng, Huttenlocher, and Newcombe (2013). One computationally-specified non-modular approach uses an associative learning model (Miller, 2009), based on a model originally formulated to explain reorientation data from non-human animals (Miller & Shettleworth, 2007) to explain findings from humans. In this account, cues compete with each other by gaining or losing strength based on a variant of the Rescorla and Wagner (1972) learning rule, adapted to encompass operant learning as well as classical conditioning. This model has the great virtue of precision in its assumptions about encoding and processing, and it provides a good fit to a variety of data. However, the extension of the original model to encompass human development relies on age-related variations in learning rate, an assumption that does not fit the developmental findings (Cheng et al., 2013). Studies with children involve very few trials (often just 4), and do not find better performance on the last trial than the first.

An alternative is to link the development of human reorientation to the development of cue combination, an idea suggested previously (e.g. Newcombe & Huttenlocher, 2006) but not computationally specified or evaluated (Cheng et al., 2013). The purpose of this paper is to specify such a computational model, and compare it with the modular encapsulation-plus-language model. We also compare it to the associative model, using the same set of data examined by Miller (2009). We evaluate the generality of its explanatory power by cross-predicting independent sets of empirical data. We restrict our scope of investigation in this paper to the development of reorientation in humans, and caution that it remains to be determined whether the model also captures the behavior of non-human species in the reorientation paradigm, for whom operant learning may be more essential (Miller & Shettleworth, 2007, 2008).

Table 1 summarizes the main differences between the proposed model, the modular hypothesis and the associative learning model of development. First, our model is grounded in the principle of cue combination as a form of probabilistic inference, suggesting that integration of information can occur from early in human development (i.e. it does not depend on language to bridge between otherwise hypothetically encapsulated modules). Second, our model does not require (although it can accommodate) a process of learning, since it is responsive to internal uncertainty based on perception and memory processes; in comparison, in the modular account, language learning is critical, as it then allows for modules to be linked, and in the associative model, learning based on external feedback is central to the reinforcement and suppression of cues. Third, our model includes a potential role of spatial language as a distinct cue that can exert an effect, but situates this effect against a background of cue combination. Thus, the inclusion of language as a strategic cue for reorientation differs from the position of modularity theory (Hermer & Spelke, 1996; Hermer-Vazquez, 1997), in that we suggest that language is not the only way that information can be combined during reorientation; rather, it acts as an independent cue that helps to reduce uncertainty in reorientation. Fourth, we use a relatively small set of cues and minimal free parameters. The associative model has a higher number of cues and adjustable parameters, due to the fact that it also parameterizes the learning process. Finally, we use a combination of fitting and cross-prediction to evaluate the models, which provides a general, rigorous way of assessing model performances. In the following sections, we show that this simple proposed cue combination model accounts for existing empirical data.

Table 1
Qualitative comparison of cue-combination and existing accounts of spatial reorientation.

Property	Modularity hypothesis	Associative model	Cue combination model
Grounding principle	Encapsulation +language	Associative learning	Probabilistic inference
Role of learning	Language only	Central	Not required for humans
Role of language	Dominant once acquired	Unaccounted for	Strategic cue
Number of free parameters	Underspecified	Relatively high	Relatively low
Method of evaluation	Empirical	Fitting	Fitting+cross prediction

The rest of this paper is organized as follows. We first illustrate the idea of cue combination in informal terms. We then present our computational model, which formalizes these ideas. We then describe the sources of empirical data on which we draw, and present three case studies in which we test our model against these empirical findings, and compare the results to those of alternative models.

2. Illustration of cue combination

Fig. 1 illustrates the overall concept of cue combination. Here and elsewhere, we assume that a person is inside a closed space (e.g. a room), has seen a target 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. Their task is to recover the target object after disorientation. For illustrative purposes, suppose that there are two independent cues in this reorientation task. Each cue provides some information about the location of the target t^*, which is fixed and located in one of four possible locations. The height of the bars for the individual cues represents the strength of each cue at each location, which varies across the two cues. The taller bars correspond to locations that a cue strongly predicts to be possible target locations. The shorter bars correspond to non-target locations. Concretely, Cue 1 is a cue based on surface geometry of the enclosure that predicts corners L_2 and L_4 as the most likely candidates for target location, because these two corners are geometrically equivalent to the actual target location (L_2). Cue 2 relies on an explicit landmark—a wall painted blue in this case—that predicts its adjacent corners L_2 and L_3 ambiguously as probable target locations, by virtue of association. Although both cues provide some degree of information in determining the location of the target, neither is sufficient to predict the target precisely. Thus, for a rational agent, an optimal strategy would be to combine information from the two cues, which would yield a substantially sharper response over the true target and hence allow for reduction of uncertainty – reflected in a resulting distribution of choice probabilities that peaks at the target location.

3. Computational formulation of cue combination

Following standard formulations of cue integration (e.g. Ernst & Banks, 2002), we model spatial reorientation as probabilistic infer-