Journal of Experimental Psychology: General
2001, Vol. 130, No. 2, 273–298

0096-3445/01/$5.00 DOI: 10.1037//0096-3445.130.2.273

Grounding Spatial Language in Perception: An Empirical and Computational Investigation

Terry Regier
University of Chicago

Laura A. Carlson
University of Notre Dame

The present paper grounds the linguistic categorization of space in aspects of visual perception; specifically, the structure of projective spatial terms such as *above* are grounded in the process of attention and in vector-sum coding of overall direction. This is formalized in the attentional vector-sum (AVS) model. This computational model accurately predicts linguistic acceptability judgments for spatial terms, under a variety of spatial configurations. In 7 experiments, the predictions of the AVS model are tested against those of 3 competing models. The results support the AVS model and disconfirm its competitors. The authors conclude that the structure of linguistic spatial categories can be partially explained in terms of independently motivated perceptual processes.

The mapping of language onto space is being actively studied in many of the disciplines of cognitive science. Examples may be drawn from neuroscience (e.g., Farah, Brunn, Wong, Wallace, & Carpenter, 1990; Shallice, 1996; Stein, 1992), from psycholinguistics (e.g. Carlson-Radvansky & Irwin, 1993, 1994; Carlson-Radvansky & Logan, 1997; Garnham, 1989; Landau & Jackendoff, 1993; Levelt, 1984; Talmy, 1983; H. A. Taylor & Tversky, 1996; van der Zee, 1996), including crosslinguistic research (e.g., Cora: Casad, 1988; Dutch: Helmantel, 1998; American Sign Language: Emmorey & Casey, 1995; Russian: Janda, 1988; Tzeltal: Brown & Levinson, 1993; Levinson, 1996; Italian: Taylor, 1988), and from computational modeling (e.g., Gapp, 1994, 1995; Harris, 1990; Herskovits, 1986, 1998; Munro, Cosic, & Tabasko, 1991; Nenov & Dyer, 1994; Regier, 1996; Schirra, 1993; Siskind, 1994; Suppes, Liang, & Boettner, 1992). Across these disciplines, researchers are asking how we link up spatial expressions with our representations of objects and their relations.

Why all this interest? One reason is that the linguistic categorization of space serves as an interface between language and the perceptual world. This makes it likely that the structure of linguistic spatial categories may be explicable in terms of perceptual processes that are not themselves linguistic in character (Hayward & Tarr, 1995; Landau & Jackendoff, 1993; see also Kay & McDaniel, 1978, for related work in the semantic domain of color). Thus, one reason to examine spatial language is that it offers the possibility of grounding some aspects of language in nonlanguage.

We bring together two of the above streams of research in spatial language, the psycholinguistic and the computational, to explore this issue of semantic grounding. Specifically, we ask what nonlinguistic perceptual processes may underlie the semantics of English projective spatial terms such as *above*. A number of researchers have examined how the acceptability of spatial terms varies across different spatial configurations (Carlson-Radvansky & Logan, 1997; Franklin, Henkel, & Zengas, 1995; Gapp, 1994, 1995; Hayward & Tarr, 1995; Logan & Sadler, 1996). For example, Logan and Sadler presented participants with a small *O* in the middle of an invisible 7 × 7 grid (in cell 4, 4). Across trials, a small *X* was placed in each of the remaining cells in the grid, and participants were asked to rate the acceptability of the sentence "The X is above the O," as a description of the spatial scene.

In general, we use the term *landmark* to refer to the central object relative to which another object is located; that other object is referred to as the *trajector* (Langacker, 1987; these are also known as the *reference* and *located* objects, respectively; Carlson-Radvansky & Irwin, 1993, 1994; Levelt, 1984). Thus, in Logan and Sadler's (1996) study, the landmark was the *O*, the trajector was the *X*, and acceptability ratings concerned the relation between the two.

These acceptability ratings were presented in a three-dimensional plot. The *x*-axis and *y*-axis represented the rows and columns, respectively, of the 7 × 7 grid, and the *z*-axis represented the mean acceptability rating, given a placement of the *X* at each

This article reflects equal contributions of both authors; order of authorship was arbitrary. This work was supported in part by National Science Foundation Grant SBR97-27638. Portions of this work were presented at the Language and Space Workshop at the 14th National Conference on Artificial Intelligence in Providence, Rhode Island, July 1997; the Workshop on Interfacing Models of Language at the Neural Information Processing Systems conference in Breckenridge, Colorado, December 1997; at the Grounding Word Meaning Workshop at the 15th National Conference on Artificial Intelligence in Madison, Wisconsin, July, 1998; at the 71st Annual Meeting of the Midwestern Psychological Association in Chicago, Illinois, May 1999; and at the 1998 meeting of the Chicago Linguistic Society in Chicago, Illinois, April 1998.

We thank participants at the conferences, as well as three anonymous reviewers, and Barbara Landau, Gordon Logan, and Annette Herskovits for helpful comments. We also thank Derrick Higgins and Bryce Corrigan for their help in programming and running the models.

Correspondence concerning this article should be addressed to Laura A. Carlson, Department of Psychology, University of Notre Dame, 118-D Haggar Hall, Notre Dame, Indiana 46556, or Terry Regier, Department of Psychology, University of Chicago, Green 414, 5848 South University Avenue, Chicago, Illinois 60637. Electronic mail may be sent to LCarlson@nd.edu.

location within the grid. Such plots, referred to as *spatial templates,* provide a visualization of the shape of the spatial category *above* relative to a particular object.

Figure 1 shows a spatial template for *above,* constructed from Logan and Sadler's (1996) data (Experiment 2). Cell (4, 4) is empty because that is where the landmark was placed. Logan and Sadler identified three regions along the surface of the spatial template: The *good* region consisted of the highest ratings and corresponded to the peak in the template. The *acceptable* region consisted of intermediate ratings and corresponded to the regions flanking the good region. The *bad* region consisted of the remaining uniformly low ratings. We refer to these regions as *direct, oblique,* and *other,* respectively, because this classification focuses on the locations of the regions rather than on the acceptability judgments commonly encountered in those regions. The overall shape of the plot is not specific to these particular landmarks and trajectors (cf. Carlson-Radvansky & Logan, 1997; Hayward & Tarr, 1995; Logan & Sadler, 1996). Nor is it specific to a particular relation (Logan & Sadler, 1996): The spatial templates for *above, below, under, over, left,* and *right* all had similar shapes and differed only in orientation.

What underlies such ratings, relative to this landmark and others? More pointedly, what perceptual or cognitive structures are reflected in these linguistic judgments? Does spatial perception shape spatial language in this instance, and if so, how? Our goal in this article is to answer these questions.

We proceed as follows. We focus on the spatial relation *above* in the expectation that our results will generalize to other projective relations that share the same spatial template shape. We consider four computational models, each one specifying a possible mechanism underlying linguistic spatial term ratings. Significantly, one of these four models is in part independently motivated by perceptual considerations. We present a set of seven experiments designed to discriminate among the models, using a variety of landmark shapes. We examine both the human behavioral data and the fits of the models to these data. To preview the findings, only the independently motivated model passes all our empirical and computational tests. It also provides a tighter overall fit to our data. This model thus provides a preliminary account of the relation between spatial perception and spatial language.

The Models

Each of the four models views a point-trajector object located relative to a two-dimensional landmark object and returns an acceptability judgment, indicating how well the spatial term *above* describes the relation between the two objects. These predicted judgments are then compared with empirically obtained data. We describe the models and then present a set of contrasting predictions that the models make. These predictions motivate our empirical and computational work.

The Bounding Box Model

The bounding box (BB) model is based on clear intuition.[1] According to this model, a trajector object is above a landmark object if it is higher than the highest point of the landmark and between its rightmost and leftmost points. Thus, this model selects the region of space projecting upward from the rectangular bounding box of the landmark object. This is illustrated in Figure 2.

The model has two conceptually distinct elements. The first element determines whether the trajector is vertically higher than the landmark, and the second determines whether it is horizontally centered with respect to the landmark. Both model elements may be implemented using the sigmoidal squashing function, as in Equation 1:

$$sig(x, gain) = \frac{1}{1 + \exp[gain \times (-x)]}. \quad (1)$$

This function returns values near 0 if x is negative and values near 1 if x is positive. More formally, $\lim_{x \to -\infty} sig(x, gain) = 0$; $\lim_{x \to \infty} sig(x, gain) = 1$; and there is a smooth interpolation in the intermediate region, with $sig(0, gain) = .5$. The value *gain* adjusts the abruptness of the transition from 0 to 1. In two-dimensional space, this function implements a *soft* half plane: $x = 0$ defines a line, and the function returns values near 1 for points on one side of that line and values near 0 for points on the other side. The gain parameter adjusts the *hardness* or *softness* of the half plane.

Given this building block, we first determine whether the trajector is higher than the landmark. The specific formulation we chose for this purpose accommodates a variety of landmark shapes. We start by defining the top of the landmark. Intuitively, the top is the set of landmark points that are exposed from above: the ones that would get wet in the rain. More formally, it is the set of landmark points (x, y) so that there exists no landmark point (x, y') with y' at a greater elevation than y. The height of the trajector with respect to the landmark is given by Equation 2:

$$height = \frac{sig(y - hightop, highgain) + sig(y - lowtop, 1)}{2}. \quad (2)$$

Here, *hightop* is the elevation of the highest point on the landmark top. Thus, $sig(y - hightop, highgain)$ defines a horizontal line touching this highest point and selects the half plane above it. Similarly, *lowtop* is the elevation of the lowest point on the landmark top. Thus, $sig(y - lowtop, 1)$ defines a horizontal line touching this point and selects the half plane above it. The value *height* is the average of these two horizontal sigmoids. The resulting value is 1 if the trajector is well above all landmark top points and 0 if it is well below all landmark top points. Intermediate positions receive intermediate values. The gain of the upper sigmoid is determined by a free parameter, *highgain,* and the gain of the lower sigmoid is fixed at 1. This gives the model some limited flexibility in fitting the data.

In this article, the horizontal line grazing the very top of the landmark is referred to as the *grazing* line. We explicitly test for its influence.

The other element of the BB model determines whether the trajector is horizontally centered. To do this, it selects the centered region: a vertical bar, bounded by the leftmost and rightmost points of the landmark, as in Equation 3:

$$center = sig(x - left, lrgain)^{lrexp} \times sig(right - x, lrgain)^{lrexp}. \quad (3)$$

[1] We thank an anonymous reviewer for suggesting the basic features of this model.

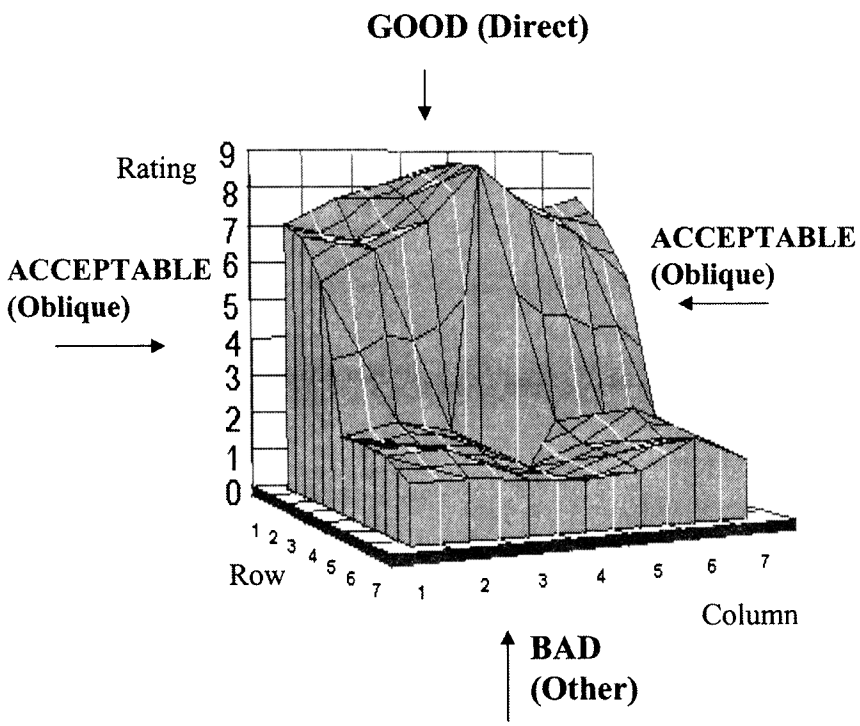

Figure 1. Spatial template for *above*. The data in the figure are from "A Computational Analysis of Apprehension of Spatial Relations," by G. D. Logan and D. D. Sadler, in *Language and Space* (p. 512) by P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), 1996, Cambridge, MA: MIT Press. Copyright 1996 by MIT Press. Adapted with permission.

Here, *sig*(x − *left, lrgain*) draws a vertical line touching the leftmost landmark point and selects the half plane to its right. Analogously, *sig*(*right* − x, *lrgain*) draws a vertical line touching the rightmost landmark point and selects the half plane to its left. The product of these two sigmoids selects the intersection of these two half planes: the centered vertical bar. There are two free parameters in this element of the model: *lrgain*, the gain on these left–right sigmoids, and *lrexp*, the power to which they are raised.

The BB model as a whole combines these two components, as in Equation 4:

$$\textit{BB model: above} = \textit{height} \times \textit{center}. \quad (4)$$

Thus, the model picks out that portion of the centered vertical bar that is higher than the landmark top. The model has three free parameters: the gain and exponent of the left- and right-flanking sigmoids and the gain of the uppermost sigmoid in the height function.[2]

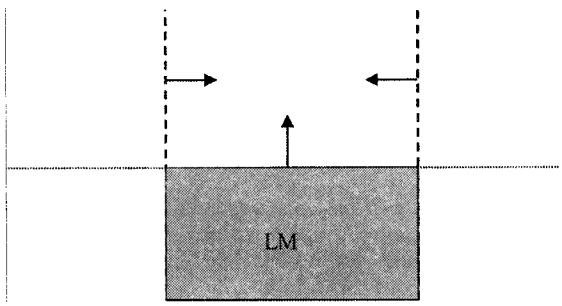

Figure 2. Illustration of the bounding box model. LM = landmark.

The Proximal and Center-of-Mass Model

The proximal and center-of-mass (PC) model assumes that projective spatial terms are defined in terms of *angular* features, that is, in polar coordinates rather than the Cartesian coordinates of the BB model. There is empirical support for such a notion. Given a point landmark and a point trajector, one may define a ray projecting from the landmark to the trajector and obtain the angular deviation of that ray from the reference axis (upright vertical for *above*). Gapp (1995) has noted that acceptability for projective spatial terms drops off roughly linearly with this angular deviation. In his data, this linearity holds for deviations from 0 through approximately 68°. Logan and Sadler's (1996) data also support the linearity of this initial drop-off, for deviations of 0 through approximately 72°. At angular deviations of 90° and beyond, acceptability ratings are essentially 0. Acceptability ratings in the intermediate range, between 72° and 90°, are not determined by this earlier work. Thus, we propose a function that smoothly interpolates in this intermediate range—the product of a line and a sigmoid, illustrated in Figure 3 and formalized in Equation 5:

[2] Further free parameters for the sigmoids of the height function were not found to be necessary.

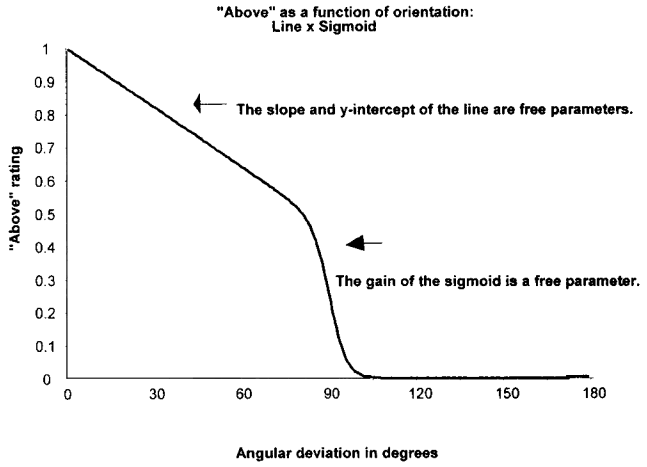

Figure 3. Alignment function of the proximal and center-of-mass model.

$$f(a) = [(slope \times a) + y\text{-}intercept] \times sig(90 - a, gain). \quad (5)$$

Here, a is the angular deviation. $f(a)$ presents a linear drop-off for $a < 90$ and then a smooth transition to 0, a reasonable formalization of these earlier empirical findings. This function has three free parameters: the slope and y-intercept of the line and the gain of the sigmoid. A central feature of this characterization of spatial term acceptability is that it is dependent only on the direction, not the length, of the vector connecting the landmark to the trajector.

This formalization assumes a point landmark. But what if the landmark object is not a point; what if it has (at least) two-dimensional extension? Regier (1996, 1997) has suggested that *above* judgments may in fact rely on two orientational features, features that are conflated in the case of a point landmark. The first feature is the *center-of-mass* orientation: the orientation of the ray connecting the center of the landmark to the center of the trajector. The second feature is the proximal orientation: the orientation of the ray connecting the edge of the landmark to the edge of the trajector, where they are closest. Regier's (1996, 1997) suggestion was that *above* judgments depend on the extent to which each of these two orientations aligns with upright vertical. Figure 4 illustrates this idea. Figure 4(a) contains a scene in which both the center of mass and proximal orientations are perfectly aligned with upright vertical; the result is a maximally acceptable example of *above*. Figure 4(b) contains a scene in which the proximal orientation is still perfectly aligned with upright vertical, but the center-of-mass orientation is not; the result is a good but less-than-perfect example of *above*. The comparison of Figures 4a and 4b thus singles out the possible role of the center-of-mass orientation in *above* judgments. Figure 4c contains a scene in which the center-of-mass orientation is the same as it was in Figure 4b but the proximal orientation is more deviant from upright vertical; the result is an even less good example of *above*. Thus, the comparison of Figures 4b and 4c suggests a role for the proximal orientation.

The PC model extends the work of Gapp (1995). As noted earlier, Gapp (1995) has found a roughly linear drop-off in acceptability with angular deviation. His data are based on the center-of-mass orientation, and the linearity is not perfect. The nature of this imperfection led Gapp to conjecture that people base their judgments not on the center-of-mass orientation, but rather on the proximal orientation. In the PC model, the idea of a linear drop-off is generalized and applied to both the proximal and center-of-mass orientations, in the expectation that both orientations are relevant. The PC model has not yet been thoroughly tested against challenging empirical data—that is one goal of this article.

Formally, the PC model characterizes *above* as a linear combination of the degrees of alignment of the center-of-mass and proximal orientations with upright vertical (Regier, 1996, 1997), as in Equation 6:

$$PC\ model\text{: } above = \alpha f(com) + (1 - \alpha) f(prox). \quad (6)$$

Here, *com* is the deviation of the center-of-mass orientation from upright vertical, *prox* is the deviation of the proximal orientation from upright vertical, and $f(\ )$ is the alignment function discussed

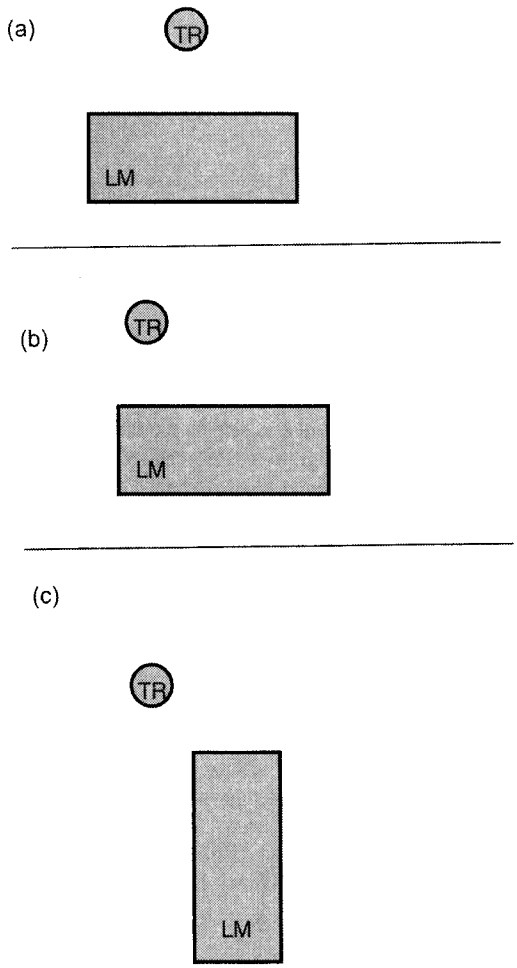

Figure 4. In panel a, both the proximal and the center-of-mass orientations are perfectly aligned with upright vertical; in panel b, the proximal orientation is still aligned with upright vertical but the center-of-mass orientation deviates somewhat; in panel c, the center-of-mass orientation retains the value it had in b, but the proximal orientation is more deviant. LM = landmark; TR = trajector.

earlier; α is a free parameter controlling the relative contribution of the center-of-mass and proximal orientations. In addition, there are three free parameters in the line × sigmoid alignment function $f(\ )$: the slope and y-intercept of the line and the gain of the sigmoid. These parameters are shared by $f(com)$ and $f(prox)$. The model as a whole thus has a total of four free parameters.[3]

The Hybrid Model

The hybrid (PC-BB) model is a variant of the PC model that also incorporates the height element of the BB model. Given a trajector located relative to a landmark, the model determines an initial acceptability judgment much as the PC model would. This initial judgment is then adjusted depending on the height of the trajector relative to the landmark. These two elements may be seen in the formal definition of the model shown in Equation 7:

$$\textit{PC-BB model: above} = [\alpha g(com) + (1 - \alpha)g(prox)] \times \textit{height}. \quad (7)$$

The first factor, $[\alpha g(com) + (1 - \alpha)g(prox)]$, is borrowed from the PC model. It is a linear combination of the degrees of alignment of the center-of-mass and proximal orientations with upright vertical. Here, *prox* and *com* are the proximal and center-of-mass orientations, $g(a)$ measures the degree of alignment of angle a with upright vertical, and α is a free parameter controlling the relative contributions of center-of-mass and proximal orientations. The orientational alignment function $g(a)$ is purely linear, as expressed in Equation 8:

$$g(a) = [(\textit{slope} \times a) + \textit{y-intercept}]. \quad (8)$$

Here, a is angular deviation from upright vertical, and *slope* and *y-intercept* are free parameters. A difference from the PC model is that there is no sigmoid gating this orientational alignment function. Instead, the model as a whole is gated by trajector height, relative to the top of the landmark.

The second model factor, height, is borrowed from the BB model and is defined in Equation 2. As before, this quantity captures the extent to which the trajector is higher than the top of the landmark. Because the PC-BB model is gated by this quantity, it is sensitive to height, as well as the proximal and center-of-mass orientations. The PC-BB model has four free parameters: α, the relative strength of the center-of-mass and proximal orientations, the *slope* and *y-intercept* of the alignment functions, and *highgain*, the gain of the uppermost sigmoid in the height function.

The Attentional Vector-Sum Model

The attentional vector-sum (AVS) model is similar in design to the PC-BB model, in that it combines an orientational component with a height function. However, its architecture is also informed by two independent observations.

The first observation is that the human apprehension of spatial relations involves *attention* (Logan, 1994, 1995). For example, Logan (1994) has found that visual search for a target in a field of distractors is slow when targets differ from distractors only in the spatial relation (either above, below, left, or right) among their elements. This suggests that spatial relations are not preattentively perceptually available, that they do not "pop out" of the visual field (Treisman & Gormican, 1988). Rather, their perception seems to require the focus of attention (see also Logan, 1995).

The second observation is that in several neural subsystems, overall direction is represented as the *vector sum* of a set of constituent directions. For example, in their studies of rhesus monkey motor cortex, Georgopoulos, Schwartz, and Kettner (1986) examined a population of orientation-tuned cells in the area of motor cortex that represents the monkey's arm. Each cell was broadly tuned, with a preferred direction, so that the cell would respond maximally when the monkey's intended arm movement was in the preferred direction. They found that the direction of motion of the arm was accurately predicted by a vector sum over the population of cells, as in Equation 9:

$$\sum_i a_i \vec{c}_i. \quad (9)$$

Here, a_i is the activation of cell i, $\vec{c}_i$ is the preferred direction of cell i, and the sum is taken over the population as a whole. Similar representations govern saccadic eye movements (Lee, Rohrer, & Sparks, 1988). In perception, Wilson and Kim (1994) have found that when participants view moving patterns, they perceive motion in the direction of the vector sum of two constituent motion components. These results suggest that the representation is a widely used one.

The AVS model brings together these two apparently unrelated observations, concerning attention and vector-sum representation. In the model, an attentional beam is focused on the landmark. In particular, the beam is focused on that point on the landmark top that is vertically aligned with the trajector or closest to being so aligned. Parts of the landmark near the center of this beam are strongly attended, whereas more distant parts of the landmark receive less attention. This yields a distribution of attention across the landmark object. In addition, at each point of the landmark, we define a vector, rooted at that position and pointing toward the trajector. This yields a population of vectors. The model then computes the vector sum over this population, weighted by attention (Moran & Desimone, 1985), following Equation 9. In this instance, however, a_i is the amount of attention currently being paid to landmark point i, and $\vec{c}_i$ is the vector rooted at landmark point i, pointing toward the trajector. The resulting vector-sum direction is then compared with upright vertical.

Figure 5 illustrates this idea. Figure 5a shows an attentional beam that has been focused on the landmark, as described above. This attentional beam radiates out to illuminate different parts of the landmark at different strengths, depending on distance from this focus. Figure 5b shows the vectors rooted at the landmark points, pointing toward the trajector. These vectors are weighted by attention at the root of each vector. The result is shown in Figure 5c: Vectors rooted nearby are more heavily weighted than distant ones. These weighted vectors are then summed. The direction of the resulting vector sum, shown in Figure 5d, is then compared with upright vertical, shown in Figure 5e.

The interaction of attention and vector summation can yield effects of the proximal and center-of-mass orientations. This is

[3] Single-factor models consisting of only one orientation feature (e.g., the proximal model and the center-of-mass model) were consistently outperformed by the PC model and hence are not presented.

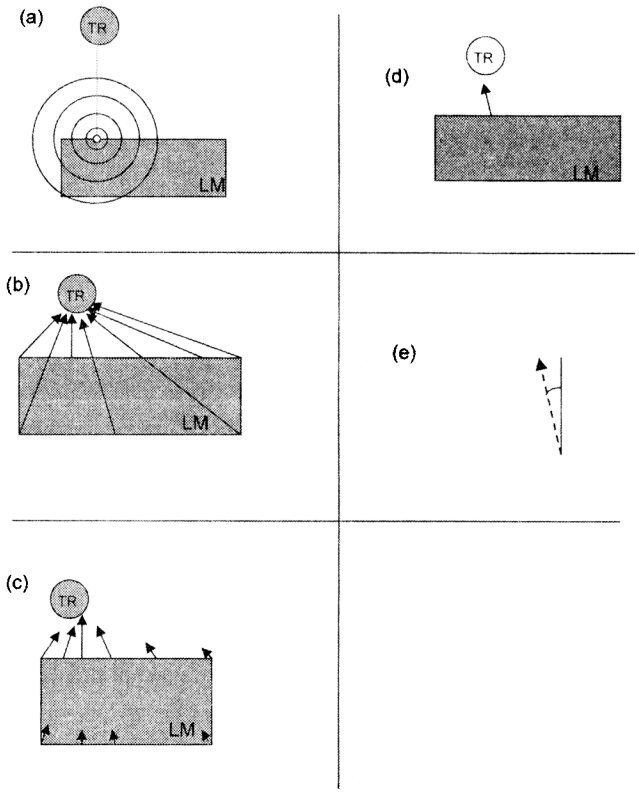

Figure 5. The attentional vector-sum model. Panel a, illustrates the attentional field, focused on the landmark (LM), near the trajector (TR). Different parts of the landmark receive different amounts of attention. Panel b illustrates the vectors rooted at each point of the landmark, pointing toward the trajector. Panel c illustrates the attentionally weighted vectors. Panel d illustrates the direction of the attentionally weighted vector sum. Panel e illustrates the orientation of the vector sum, relative to vertical upright.

seen most easily by considering what happens as the attentional field varies in width.

Narrow beam. At one extreme, the attentional field is so narrow that only the landmark point at the center of the beam receives any attention. This will be that point on the landmark top that is vertically aligned with the trajector or closest to being so aligned. This point often coincides with the landmark point that is nearest the trajector in Euclidean distance, as in Figure 5. The orientation of the vector rooted at that point is, by definition, the proximal orientation. Thus, under such circumstances, the vector sum reduces to the proximal orientation, and the model's responses are based on that feature.

Wide beam. At the other extreme, the attentional field is infinitely wide, exhibiting no drop-off in attention with distance. Under such circumstances, the vector sum always yields the center-of-mass orientation, as shown in Appendix A. Thus, with a uniform attentional field, the AVS model reduces to one that bases its predictions on the center-of-mass orientation.

Attentional field widths between these two extremes can therefore produce vector-sum directions that are intermediate between the center-of-mass orientation and the proximal orientation. The model's responses will then exhibit influences of both the center-of-mass and proximal orientations. This dependence of model behavior on attentional width implicitly raises a question: What attentional drop-off function do humans actually possess? In their empirical studies of spatial attention, Downing and Pinker (1985) have found an attentional drop-off that is roughly exponential in distance (see also LaBerge & Brown, 1989). Thus, we use an exponential decay function in the AVS model. This is intermediate between the two extremes outlined above; thus, we expect both proximal and center-of-mass effects. The specific function is shown in Equation 10:

$$a_i = \exp[-(\text{Euclidean distance from focus to point } i \text{ of landmark})/(\lambda\sigma)]. \quad (10)$$

Here, the steepness of the exponential drop-off is governed by two values, λ and σ; λ is a free parameter, and σ is the Euclidean distance between the trajector and the focus point, that is, the landmark point on which the attentional beam is centered. Thus, the farther the trajector is from the landmark, the wider the attentional beam will be. This is natural, because the perceiver must encompass both the trajector and a part of the landmark within the attentional beam.[4] Thus, a nearby trajector requires only a fairly narrow attentional beam, but a distant trajector requires a wider one.

The AVS model as a whole determines the alignment of the attentional vector sum with upright vertical and then gates the resulting quantity by trajector height. The form of the complete model is shown in Equation 11:

$$\textit{AVS model: above} = g\Big(\sum_i a_i\vec{c}_i\Big) \times \textit{height}. \quad (11)$$

Here, $g(\,)$ is the angular alignment function used in the PC-BB hybrid model. The value *height* is the height function used in both the BB and PC-BB models. This quantity indicates whether the trajector is strictly higher than (1.0), strictly lower than (0.0), or at the same height as (intermediate values) the top of the landmark. The model has four free parameters: the slope and y-intercept of the linear alignment function; λ, which controls the width of the attentional field; and the gain of the upper sigmoid in the height function.

Contrasting Predictions

These models make a number of contrasting predictions. Some predictions are shared by more than one model, whereas others are unique. For each prediction, we indicate which models give rise to it and why.

Proximal and center-of-mass orientations. The PC, PC-BB, and AVS models all predict effects of the proximal and center-of-mass orientations. Specifically, acceptability ratings should drop off with deviation of either orientation from upright vertical. The PC and PC-BB models predict this because they are explicitly based on these features. In the AVS model, these effects emerge from the attentional vector-sum operation. The BB model does not make these predictions, because it is not orientation based. Instead,

[4] We assume that the trajector itself must receive a fair amount of attention; otherwise its location would be unclear.

it predicts that acceptability ratings will drop off with horizontal or vertical distance from the lines defining the bounding box. These predictions are tested in Experiments 1–4.

Grazing line. The BB, PC-BB, and AVS models all predict that the horizontal line grazing the top of the landmark will affect acceptability ratings. Specifically, a point just above this line should receive a higher acceptability rating than a corresponding point just below it. This is predicted because these models explicitly gate their acceptability ratings by the grazing line. In contrast, the PC model contains no grazing line, and therefore does not make this prediction. These predictions are tested in Experiments 5–6.

Distance. Imagine holding a small marble half an inch above a large book lying on a table and moving the marble around over the surface of the book but maintaining the half-inch height. The AVS model predicts that at low heights such as this, *above* ratings will be relatively insensitive to the extent to which the marble is centered above the book. However, ratings should be more sensitive to centeredness if the marble is much higher above the book. Why? Because at low heights, the attentional field will be narrow, and not much of the book will receive appreciable amounts of attention—and critically, the edges of the book will not receive much attention. This means that for all intents and purposes, the marble is located above a limitless plane. When the marble is higher above the book, the attentional beam grows larger and begins to take in the edges of the book, not just the surface, and this will lead to a greater sensitivity to centeredness. The PC and PC-BB models make exactly the opposite prediction. In both of these models, acceptability decreases with deviation of the center-of-mass orientation from upright vertical. This deviation covers a greater range in the case of the low trajector than in the case of the high one. Therefore, there should be greater sensitivity to centeredness in the low case. The BB model does not predict a difference between the two cases. These predictions are tested in Experiment 7.

These three predictions guide our experimental and computational work and allow us to discriminate among the models.

Initial Model Fits

As an initial test of model feasibility, we evaluated each model's fit to the *above* data of Logan and Sadler (1996, Experiment 2). Once the models were fit to these data, the resulting parameter settings were saved and used in testing the models on our own data. This was done for two reasons. The first reason was to ensure compatibility of data across researchers; a model that fit data from two laboratories was deemed stronger than a model that only fit data from one. The second reason was to reward models that generalized well to novel landmark shapes. The stimuli in the Logan and Sadler (henceforth LS) data set consist of a very small trajector located relative to a very small landmark. Our stimuli, in contrast, contain landmarks with appreciable two-dimensional extension. Thus, if a model generalizes well from the LS data to ours, that will suggest that it captures more than just the particulars of the LS data set: It might also capture the general processes underlying spatial language, as applied to a variety of landmarks.

The models were fit to the LS data by an inverse hill-climbing procedure, inverse in that it performs descent rather than ascent in its objective function. The algorithm incrementally adjusts the parameters of the model so as to minimize the error—the discrepancy between the empirically observed and model-predicted values. The procedure terminates when no further parameter changes yield a reduction in error. In all cases, the step size for parameter adjustment was 0.001. Parameters were initialized as follows: 1.0 for sigmoid exponents (BB model), alignment function y-intercepts (all other models), top sigmoid gains (BB, PC-BB, AVS), and λ, controlling the width of the attentional field (AVS); 0.5 for α, the relative weight of the proximal and center-of-mass orientations (PC, PC-BB); and 0.0 for all remaining sigmoid gains (BB, PC) and alignment function slopes (PC, PC-BB, AVS). These initialization values led to tight fits for all models. No tighter fits on the LS data were obtained with a wide range of other initialization values.

The resulting parameter settings are shown in Table 1. Linear regression was performed to measure how well the models' output predicted the LS ratings, given these parameter settings. Table 2 shows the results of this exercise. For each model, the table presents R^2, a measure of the goodness of fit, with a value of 1.0 indicating a perfect fit. The table also presents adjusted R^2, an adjustment of the original R^2 value that supports comparisons among models with different numbers of free parameters. Also shown are the slope and y-intercept of the regression lines for each model.

All four models fit the LS data fairly well, particularly the PC, PC-BB, and AVS models. This established the initial feasibility of all models. To more effectively distinguish among the models, we tested the contrasting predictions in a series of experiments. The experiments are grouped in three sections, each testing one of the three major predictions made by the models. Thus, Experiments

Table 1
Parameter Settings for Each Model

Model parameter	Parameter value: Logan & Sadler (1996)	Parameter value: Experiment 7
BB		
Gain on left–right sigmoids	0.109	0.065
Gain on top sigmoid	0.066	0.373
Exponent for left–right sigmoids	0.062	0.220
PC		
α, relative weight of P and C	0.500	0.174
y-intercept of alignment function	0.969	0.929
Slope of alignment function	−0.005	−0.006
Gain on sigmoid	0.112	3.265
PC-BB		
α, relative weight of P and C	0.500	0.115
y-intercept of alignment function	1.007	0.909
Slope of alignment function	−0.006	−0.005
Gain on top sigmoid	0.131	6.114
AVS		
λ, attentional field width	1.000	0.512
y-intercept of alignment function	1.007	1.224
Slope of alignment function	−0.006	−0.007
Gain on top sigmoid	0.131	0.002

Note. Parameter values were from Logan and Sadler's (1996, Experiment 2) *above* data and from our Experiment 7. The Logan and Sadler parameter values were used for all simulations unless otherwise noted in text. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum.

Table 2
Model Fits to Above Data From Logan and Sadler (1996)

Model	R^2	Adj R^2	Slope	*y*-intercept
BB	.904	.897	0.907	0.038
PC	.959	.955	1.011	−0.024
PC-BB	.963	.959	1.030	−0.036
AVS	.963	.959	1.030	−0.036

Note. Adj = adjusted; slope and *y*-intercept are from regression lines of models. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum.

1–4 test for the proximal and center-of-mass orientations; Experiments 5 and 6 test for the grazing line; and Experiment 7 tests for an effect of distance. Within each section, we also determine how well the models fit the data collected for each experiment and whether the models exhibit the qualitative effects that each experiment was designed to test.

Testing for Proximal and Center-of-Mass Orientation

Experiments 1–2

The PC, PC-BB, and AVS models predict that acceptability ratings will exhibit effects of the proximal and center-of-mass orientations. The BB model does not predict these effects. To test for independent effects of proximal and center-of-mass orientation, in Experiment 1, we manipulated proximal orientation while holding center-of-mass orientation constant, whereas in Experiment 2, we manipulated center-of-mass orientation while holding proximal orientation constant. In both experiments, we used horizontally extended (i.e., wide) and vertically extended (i.e., tall) rectangles as landmarks and smaller squares as trajectors. Landmarks appeared centered in the middle of an invisible 5 × 5 grid, and across trials, the trajector was placed in each of the empty cells of the grid. Following Logan and Sadler (1996), participants rated the acceptability of the vertical relations (*above* and *below*) and horizontal relations (*left* and *right*). The only difference between Experiments 1 and 2 was in the placement of the trajectors.

In Experiment 1, trajectors were always placed in the same location within a cell, so that the center-of-mass orientation was equated across the tall and wide landmarks. This necessarily resulted in different proximal orientations for the tall and wide landmarks within the oblique region. Table 3 gives the proximal orientations for vertical and horizontal relations for tall and wide landmarks. If proximal orientation influences spatial term acceptability in the manner suggested by some models, then acceptability would decline as this orientation deviated from the cardinal axis (upright vertical for *above*). For vertical relations, the proximal orientations for the tall landmark were larger than they were for the wide landmark; however, for horizontal relations, the proximal orientations for the wide landmark were larger than for the tall landmark. Therefore, an influence of proximal orientation would be observed as an interaction between spatial relation (vertical vs. horizontal) and landmark (tall vs. wide).

In Experiment 2, trajectors were placed in different locations within a cell for the tall and wide landmarks, thereby equating proximal orientation and necessarily varying center-of-mass orientation. Table 4 gives the center-of-mass orientation for vertical and horizontal relations for tall and wide landmarks. If center-of-mass orientation influences the acceptability of a spatial relation, then the greater the orientation, the less acceptable the spatial relation would be. For vertical relations, the center-of-mass orientation for the wide landmark was larger than the center-of-mass orientation for the tall landmark; however, for horizontal relations, the center-of-mass orientation for the tall landmark was larger than that for the wide landmark. Therefore, an influence of center-of-mass orientation would be observed as an interaction between spatial relation (vertical vs. horizontal) and landmark (tall vs. wide).

Table 3
Proximal Orientations for Tall and Wide Landmarks for a Vertical Relation and a Horizontal Relation for Experiment 1

	Tall					Wide				
Relation	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
					Vertical					
R1	49.5	27.5	0	27.5	49.5	40.5	16.8	0	16.8	40.5
R2	73.2	56	0	56	73.2	62.5	34	0	34	62.5
R3	90	90	[box]	90	90	90	90	[box]	90	90
R4	106.8	124	180	124	106.8	117.5	146	180	146	117.5
R5	130.5	152.5	180	152.5	130.5	139.5	163.2	180	163.2	139.5
					Horizontal					
R1	40.5	62.5	90	117.5	139.5	49.5	73.2	90	106.8	130.5
R2	16.8	34	90	146	163.2	27.5	56	90	124	152.5
R3	0	0	[box]	180	180	0	0	[box]	180	180
R4	16.8	34	90	146	163.2	27.5	56	90	124	152.5
R5	40.5	62.5	90	117.5	139.5	49.5	73.2	90	106.8	130.5

Note. The center boxes represent the reference objects. R1–R5 refer to the rows and C1–C5 refer to the columns of the 5 × 5 grid that defined placement of the located object.

Table 4
Center-of-Mass Orientations for Tall and Wide Landmarks for a Vertical Relation and a Horizontal Relation

	Tall					Wide				
Relation	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
					Vertical					
R1	40.2	23.8	0	23.8	40.2	49.8	35.2	0	35.2	49.8
R2	54.8	36.5	0	36.5	54.8	66.2	53.5	0	53.5	66.2
R3	90	90	—	90	90	90	90	—	90	90
R4	125.2	143.5	180	143.5	125.2	113.8	126.5	180	126.5	113.8
R5	139.8	156.2	180	156.2	139.8	130.2	144.8	180	144.8	130.2
					Horizontal					
R1	49.8	66.2	90	113.8	130.2	40.2	54.8	90	125.2	139.8
R2	35.2	53.5	90	126.5	144.8	23.8	36.5	90	143.5	156.2
R3	0	0	—	180	180	0	0	—	180	180
R4	35.2	53.5	90	126.5	144.8	23.8	36.5	90	143.5	156.2
R5	49.8	66.2	90	113.8	130.2	40.2	54.8	90	125.2	139.8

Note. The center boxes represent the reference objects. R1–R5 refer to the rows and C1–C5 refer to the columns of the 5 × 5 grid that defined placement of the located object.

Method

Participants. Thirty-six University of Notre Dame undergraduates participated in Experiment 1, and 39 participated in Experiment 2. All participants in these and subsequent experiments gave informed consent. Participants were compensated with partial credit in an undergraduate psychology course.

Stimuli. The displays were constructed on the basis of a 5 × 5 matrix that was invisible to participants. The matrix measured 95 mm along a side, and each cell within the matrix measured 19 mm along a side. Displays were presented on a computer monitor; viewing distance was unconstrained, but was approximately 51 cm. The landmark was always placed in the middle of the matrix (i.e., in cell 3, 3). The landmark was either a tall red rectangle (5 mm wide × 15 mm tall) or a wide green rectangle (15 mm wide × 5 mm tall).[5]

In addition to the landmark, each picture display contained a trajector, which was a 3-mm blue square. In Experiment 1, the trajector was placed in the same location within each cell of a 5 × 5 grid surrounding the landmark, thereby equating center-of-mass orientation across the tall and wide landmarks but varying proximal orientation. In Experiment 2, the position of the square within each cell of the matrix was varied (i.e., it was not always centered in the cell) across the tall and wide rectangles, thereby equating proximal orientation but varying center-of-mass orientation. Pixel locations for all trajectors and landmarks for Experiments 1–7 and empirical and model-predicted ratings for all trajector placements in these experiments are provided at http://www.ccp.uchicago.edu/~regier/avs.

Before each picture display, a sentence appeared of the form "The square is _____ the red/green block." The blank was filled in with one of the following spatial relations: "above," "below," "to the left of," or "to the right of." The spatial relation was capitalized and centered for emphasis. Accompanying the sentence was a reminder of the acceptability rating scale (see below) and an instruction to press the return key to see the picture display.

Procedure. Participants were told that they would be rating the acceptability of sentences as descriptions of pictures using a 10-point scale, ranging from 0 (*not at all acceptable*) to 9 (*perfectly acceptable*). They were encouraged to use intermediate values if they wished. Each participant completed 5 practice trials, followed by 480 experimental trials (5 landmarks × 4 spatial relations × 24 unoccupied matrix locations). The trials were self-paced and were presented in a unique random order for each participant. Each experiment took about 50 min to complete.

Results and Discussion

Empirical data: Experiment 1. Mean acceptability ratings for all placements of the trajector for each spatial relation are provided in Appendix B. An influence of proximal orientation would result in the following pattern: For vertical relations (averaging across *above* and *below*), wide landmarks would receive higher acceptability ratings than the tall landmarks because the proximal orientation for the tall landmarks was greater than the proximal orientation for the wide landmarks; however, for horizontal relations (averaging across *left* and *right*), tall landmarks would receive higher acceptability ratings than wide landmarks because the proximal orientation for wide landmarks was greater than the proximal orientation for tall objects. To test these predictions, a 2 (spatial relation: vertical vs. horizontal) × 2 (landmark shape: tall alone vs. wide alone) repeated measures analysis of variance (ANOVA) was conducted on the mean acceptability ratings calculated across cells in the oblique region.[6] Unless otherwise noted, the p value adopted for significance for this and all other analyses was .05.

There was a main effect of spatial relation, $F(1, 35) = 5.0$, $p < .032$, $MSE = 0.33$, with vertical relations ($M = 6.6$) rated signif-

[5] Participants also gave ratings to three other landmarks, each one made from superimposing the tall and wide rectangles. On some trials, participants were asked to interpret the relation with respect to the wide block; on some, with respect to the tall block; and on some, with respect to both blocks. Data for trials with reference to the tall and wide blocks replicate the effects reported when the wide and tall landmarks were presented individually. Because the individual tall and wide landmarks were the test cases for the models, only these data are reported.

[6] Similar analyses were also conducted on the mean acceptability ratings for the direct and other regions for both Experiments 1 and 2. All Fs < 1.

icantly higher than horizontal relations (M = 6.4). Superior performance for vertical terms relative to horizontal terms has been consistently found and has been attributed to the fact that vertical relations are aligned with salient perceptual and environmental cues (e.g., gravity, the top and bottom of the display) whereas left and right are not explicitly marked (e.g., Clark, 1973; Franklin & Tversky, 1990; Levelt, 1984). There was no effect of landmark shape ($p > .12$). Most critically, there was a significant interaction between spatial relation and landmark shape, $F(1, 35) = 16.8$, $p < .001$, $MSE = 0.13$. This interaction took the expected form: For vertical relations, wide objects (M = 6.7) received higher ratings than tall objects (Ms = 6.5); however, for horizontal relations, tall objects (M = 6.6) received higher ratings than wide objects (Ms = 6.3). A 95% confidence interval constructed on the basis of the error term for the interaction revealed a critical difference of .17 for significance; thus, the tall versus wide comparisons were significant for both vertical and horizontal relations. These data indicate that proximal orientation significantly influenced the parsing of space around an object when center-of-mass orientation was held constant.

Empirical data: Experiment 2. Mean acceptability ratings for all placements of the trajector for each spatial relation are provided in Appendix C. An influence of center-of-mass orientation would result in the following pattern: For vertical relations, tall landmarks would receive higher ratings than wide landmarks because the center-of-mass orientation was greater for wide than for tall landmarks; for horizontal relations, wide landmarks would receive higher ratings than tall landmarks because the center-of-mass orientation was greater for tall than for wide. To test this prediction, a 2 (spatial relation: vertical vs. horizontal) × 2 (landmark shape: tall vs. wide) repeated measures ANOVA was conducted on the mean acceptability ratings calculated across cells in the oblique region. There was a main effect of spatial relation, $F(1, 38) = 9.2$, $p < .004$, $MSE = 0.17$, with vertical relations (M = 6.7) rated significantly higher than horizontal relations (M = 6.5). There was no effect of landmark shape ($F < 1$). Most important, there was a significant interaction between spatial relation and landmark shape, $F(1, 38) = 12.4$, $p < .001$, $MSE = 0.18$. This interaction took the expected form: For vertical relations, tall landmarks (M = 6.8) received higher ratings than wide landmarks (M = 6.5); however, for horizontal relations, wide landmarks (M = 6.6) received higher ratings than tall landmarks (M = 6.3). A 95% confidence interval constructed on the basis of the error term for the interaction revealed a critical difference of .20 for significance. Thus, the tall versus wide comparisons were significant for both vertical and horizontal relations. These data indicate a significant influence of center-of-mass orientation when proximal orientation was held constant. Taken together, Experiments 1 and 2 provide empirical verification of the importance of the center-of-mass and proximal orientations in spatial term ratings, as predicted by the PC, PC-BB, and AVS models.

Model fits and simulations. To further probe the models, we tested them on the *above* data collected for Experiments 1 and 2. We were interested in two issues. First, we investigated, quantitatively, how tightly the models fit the data. Second, we determined whether each model exhibited qualitative effects of the proximal and center-of-mass orientations. This was determined using the logic of the experiments, applied to model outputs rather than human responses.

We retained the parameter settings obtained from fitting Logan and Sadler's (1996) data and measured the models' generalization to the new data. This was done by presenting the models with our experimental stimuli, recording the model's output, and determining through regression how well the model output predicted the empirically obtained acceptability rating. This was done separately for the two landmarks within each experiment, giving us a total of four data sets for Experiments 1–2. The results are presented in Table 5. This table shows R^2, adjusted R^2, and the slope and y-intercept of the regression line. All four models performed well on the data collected from these experiments. The PC-BB and AVS models in particular provided very good fits.

To determine whether the models exhibited qualitative effects of the proximal and center-of-mass orientations, we presented the critical stimuli for the various experiments to the models and recorded the output. By the logic of Experiment 1, an effect of the proximal orientation would be reflected in *above* ratings in the oblique regions. Specifically, these model-predicted ratings would be higher for the wide landmark than for the tall landmark. Therefore, given each model's predicted ratings, we measured the average oblique rating relative to the wide landmark, minus the average oblique rating relative to the tall landmark. If this number was positive, the model would show an effect of the proximal orientation. By the logic of Experiment 2, an effect of the center-of-mass orientation would be reflected in higher *above* ratings for the tall landmark, as compared with the wide landmark. Therefore, we measured the average oblique rating relative to the tall landmark minus the average oblique rating relative to the wide landmark. If this number was positive, the model would show an effect of the center-of-mass orientation.

Table 6 contains the results of these tests. All four models exhibited an effect of the proximal orientation, and all models except for BB exhibited an effect of the center-of-mass orientation. It is clear why the PC, PC-BB, and AVS models exhibited these effects: They predicted them from the outset.[7] The BB model exhibited an effect of proximal orientation, but not center-of-mass orientation, due to the specific nature of the trajector placements. In Experiment 1, the placements were arranged so as to hold the center-of-mass orientation constant while manipulating the proximal orientation. However, these manipulations also affected the

Given that proximal and center-of-mass models did not vary across these regions, the expected interaction between relation and landmark shape was not significant ($F < 1$), except for the analysis of the direct region in Experiment 2, $F(1, 38) = 6.0$, $p < .019$, $MSE = 0.4$. The form of this interaction was directly opposite from the pattern of data in the oblique region. Specifically, vertical relations with wide landmarks (M = 8.9) were rated higher than with tall landmarks (M = 8.5), whereas horizontal relations with tall landmarks (M = 8.9) were rated higher than with wide landmarks (M = 8.8).

[7] The AVS model predicts an effect of the proximal orientation under particular conditions. The focus of the attentional beam at the landmark top point that is vertically aligned with the trajector or closest to being so aligned must coincide with the landmark point that is closest to the trajector in euclidean distance. This was the case for the trajector placements examined in this experiment.

Table 5
Model Fits to Data from Experiments 1–7, Broken Down by Landmark Shape

Model	R^2	Adj R^2	Slope	*y*-intercept
Experiment 1				
Tall rectangle (24 data points)				
BB	.982	.979	0.945	−0.387
PC	.963	.955	0.975	0.073
PC-BB	.995	.994	1.075	−0.616
AVS	.996	.995	1.088	−0.614
Wide rectangle (24 data points)				
BB	.971	.966	0.981	−0.076
PC	.954	.944	0.989	0.202
PC-BB	.993	.992	1.075	−0.325
AVS	.994	.993	1.060	−0.323
Experiment 2				
Tall rectangle (24 data points)				
BB	.995	.994	0.998	−0.391
PC	.973	.968	0.998	0.007
PC-BB	.992	.991	1.085	−0.635
AVS	.993	.992	1.098	−0.637
Wide rectangle (24 data points)				
BB	.984	.981	1.010	−0.372
PC	.960	.952	0.991	−0.173
PC-BB	.994	.993	1.070	−0.728
AVS	.995	.993	1.056	−0.721
Experiment 3				
Tall rectangle (56 data points)				
BB	.979	.977	1.013	−0.511
PC	.883	.873	0.903	0.936
PC-BB	.983	.982	1.079	−0.626
AVS	.984	.983	1.060	−0.596
Wide rectangle (56 data points)				
BB	.988	.987	1.062	−0.310
PC	.968	.965	0.937	0.187
PC-BB	.992	.992	1.024	−0.431
AVS	.993	.992	1.017	−0.407
Experiment 4				
Upright triangle (4 data points)				
BB	.963		1.048	0.212
PC	.967		0.697	2.373
PC-BB	.993		1.485	−3.723
AVS	.991		1.402	−2.859
Inverted triangle (4 data points)				
BB	.999		1.102	−0.329
PC	.987		1.037	−0.161
PC-BB	.986		1.150	−0.909
AVS	.990		1.150	−0.907
Experiment 5				
L shape (65 data points)				
BB	.943	.941	0.960	−0.141
PC	.862	.853	0.851	0.546
PC-BB	.943	.940	0.870	−0.417
AVS	.976	.975	0.944	−0.462
Experiment 6				
Tall triangle (31 data points)				
BB	.750	.723	0.921	0.665
PC	.770	.734	0.995	0.976
PC-BB	.898	.882	1.009	0.270
AVS	.930	.919	1.168	0.169
Experiment 7				
Wide rectangle (25 data points)				
BB	.932	.922	1.070	−0.567
PC	.957	.949	1.015	−0.220
PC-BB	.916	.900	1.043	−0.819
AVS	.965	.958	1.068	−0.836
Critical points only (14 data points)				
BB	.097		0.925	0.389
PC	.407	.143	1.174	−1.666
PC-BB	.173		1.496	−4.726
AVS	.555	.357	1.820	−7.166
Fit directly to data (25 data points)				
BB	.983	.981	1.130	−0.948
PC	.978	.974	1.054	−0.447
PC-BB	.978	.974	1.102	−0.846
AVS	.985	.982	1.028	−0.236
Fit directly to data, critical points only (14 data points)				
BB	.784	.719	1.635	−5.103
PC	.400	.133	0.428	4.552
PC-BB	.367	.086	0.243	6.015
AVS	.888	.838	1.138	−1.287
Composite (337 data points)				
BB	.953	.952	1.007	−0.242
PC	.910	.909	0.926	0.443
PC-BB	.959	.958	1.01	−0.450
AVS	.970	.970	1.031	−0.439

Note. Adj = adjusted; slope and *y*-intercept are from regression lines of model. Unless noted otherwise, model parameter settings were obtained by fitting the models to Logan and Sadler's (1996) data. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum.

vertical and horizontal distance between the trajector and the landmark. For the wide landmark, vertical distance was greater and horizontal distance was less than for the tall landmark. Both of these factors contributed to higher ratings for the wide landmark. Thus, for these placements, the BB model produced results that matched our empirical findings.

However, the BB model's failure to exhibit a center-of-mass effect in Experiment 2 argued against this model. In this experiment, the trajector placements were arranged so as to hold the proximal orientation constant while manipulating the center-of-mass orientation. These trajector placements also held constant both the vertical and the horizontal distance to the nearest edge of the landmark. Because these BB-relevant features did not vary within Experiment 2, the model did not exhibit a center-of-mass effect. (In fact, the BB model exhibited a small effect in the opposite direction. This was the influence of the sigmoid that marked the far edge of the landmark, opposite the trajector. Because the trajector was farther from this sigmoid in the case of the wide landmark, the model produced a slightly higher rating for this landmark.) A center-of-mass effect was found empirically, with the same stimuli. This disconfirmed the BB model. Critically, this disconfirmation holds not just for these parameter settings of the

Table 6
Effects of Proximal and Center-of-Mass Orientations

Experiment	Model				Empirical data
	BB	PC	PC-BB	AVS	
1	0.596	0.347	0.361	0.092	0.093
2	−0.003	0.329	0.333	0.583	0.114

Note. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum. Experiment 1 tested for the proximal orientation; Experiment 2 tested for the center-of-mass orientation. A positive value indicates that the model exhibited the effect in question.

BB model, nor just for the particular implementation of the BB model tested here. Rather, it holds for any parameter setting of any implementation of the model, as long as the model is fundamentally based on vertical and horizontal distance.

Experiment 3

In Experiments 1 and 2, multiple locations within the oblique region were probed to assess the influence of proximal and center-of-mass orientation. However, only two locations were probed within the direct region, and at those locations the proximal and center-of-mass orientations were confounded. Thus, the experiments did not offer an opportunity to examine the effects of orientation within the direct region. This limitation also held for Logan and Sadler's (1996) spatial templates, as well as the templates constructed in other studies (e.g., Carlson-Radvansky & Logan, 1997; Hayward & Tarr, 1995). Thus, one purpose of Experiment 3 was to examine the possibility of orientation effects within the direct region. A second purpose was to replicate the center-of-mass effect obtained in Experiment 2. This was important given that the only other researcher to examine orientation effects on spatial term acceptability seemed to suggest that only proximal orientation played a significant role (Gapp, 1995). Finally, it was not clear that the exact location of the center of mass was relevant, particularly for highly elongated landmark objects, for example, a lamp above a long conference table.[8] To examine this, in Experiment 3, we used a larger wide rectangular landmark that permitted various placements within its direct region. This varied center-of-mass orientation while holding proximal orientation constant.

Method

Participants. Thirty-six University of Notre Dame undergraduates participated in exchange for partial course credit in an undergraduate psychology class. None had participated in Experiments 1 or 2.

Stimuli. Two landmarks were used: a tall rectangle and a wide rectangle. These landmarks (25 mm × 75 mm) were five times the size (in both dimensions) of the landmarks used in Experiments 1 and 2 and permitted multiple placements of the trajector along the surfaces of the landmark (see Table 7). In all, the trajector was placed in 56 locations around each landmark. The trajector was a circle with a radius of 1 pixel (0.42 mm). As in Experiments 1 and 2, a text display preceded the onset of the picture. This display contained the rating scale, a sentence of the form "The circle is above/below/to the left of/to the right of the rectangle," and instructions *to press the return key to see the picture.*

Procedure. The procedure from Experiments 1 and 2 was used. Participants rated the acceptability of a sentence as a description of a picture containing a landmark and a trajector. In all, participants performed 448 trials (2 landmarks × 56 placements × 4 relations).

Results and Discussion

Empirical data. The possible influence of the center-of-mass orientation can be assessed by examining acceptability ratings for a single spatial term. Accordingly, in the interest of space we focus on the data for *above* while noting that similar effects hold for the other relations. This is reasonable given that the relations *above, below, left,* and *right* all belong to the same class of relations (Logan & Sadler, 1996), as illustrated by the similarity across the three regions in the spatial templates for these relations in Experiments 1 and 2.

Mean acceptability ratings for *above* for placements of the trajector around the wide and tall landmarks are given in Table 7. For the wide rectangle, the cells of interest for assessing contributions of center-of-mass orientation are in rows 1–2, columns 3–9. Mean acceptability ratings for these cells were regressed on the absolute value of the center-of-mass orientations using linear regression. The regression equation revealed a significant influence of center of mass (beta weight = −.78, $p < .001$, overall $R^2 = .61$). This indicates that as the center-of-mass orientation becomes more deviant from upright, the relation becomes less acceptable.

A similar pattern of results was found with the tall rectangle. The cells of interest are in rows 1–2, columns 5–7. Mean acceptability ratings for these cells were regressed on the absolute value of the center-of-mass orientations. There was a significant influence of center of mass (beta weight = −.96, $p < .004$, overall $R^2 = .92$). These results further support the findings from Experiment 2 in showing a significant influence of the center-of-mass orientation when proximal orientation was held constant, contrary to Gapp (1995). Moreover, this influence was found in multiple regions across the spatial template: in the direct region in the current experiment and in the oblique region in Experiment 2.

Model fits and simulations. The results of the models' fits to the data are shown in Table 5. Each model was tested for a center-of-mass effect using the stimuli of Experiment 3 and examining the model's predictions for the points in the top row of the direct region above the wide landmark (points 3–9, in columns 3–9 in Table 7). We used the wide landmark because its width allowed more opportunity for a center-of-mass effect to reveal itself. The results are displayed in Figure 6.

The empirical data are displayed in Figure 6A, and the predictions of the different models are shown in Figure 6B. The empirical data show a clear peak at the center of the landmark. All four models also show such a peak. When compared with the empirical data, the peaks of the PC, PC-BB, and AVS models are somewhat too sharp, and the peak of the (already disconfirmed) BB model is somewhat too flat. But all models do exhibit a peak and therefore pass this qualitative test. Thus, for the model competition, this experiment did not eliminate any further models. The remaining models were still the PC, PC-BB, and AVS models.

[8] We thank Gordon Logan for this observation.

Table 7
Mean Ratings for Above by Landmark and Placement for Experiment 3

	Column										
Row	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
						Wide Rectangle					
R1	6.9	6.7	**8.0**	**7.9**	**8.7**	**8.9**	**8.7**	**8.5**	**7.4**	7.2	6.9
R2	5.9	6.7	**8.0**	**8.3**	**8.0**	**8.8**	**8.7**	**8.3**	**7.4**	6.4	6.4
R3	2.7	3.2	—	—	—	—	—	—	—	2.9	2.7
R4	0.8	0.9	—	—	—	—	—	—	—	0.6	1.3
R5	0.5	0.2	—	—	—	—	—	—	—	0.4	0.3
R6	0.2	0.2	0.3	0.3	0.6	0.5	0.3	0.2	0.1	0.2	0.3
R7	0.1	0.1	0.3	0.3	0.5	0.4	0.5	0.2	0.5	1.0	0.2
						Tall Rectangle					
R1			7.2	7.5	**8.4**	**8.9**	**8.2**	7.7	7.0		
R2			6.7	6.5	**8.4**	**8.9**	**8.3**	7.2	7.1		
R3			3.0	3.0	—	—	—	3.1	4.1		
R4			1.6	1.4	—	—	—	1.5	1.4		
R5			1.0	1.3	—	—	—	1.1	1.4		
R6			0.9	0.6	—	—	—	0.9	0.7		
R7			0.5	0.2	—	—	—	0.5	0.6		
R8			0.5	0.4	—	—	—	0.6	0.2		
R9			0.3	0.2	—	—	—	0.3	0.6		
R10			0.2	0.2	0.1	0.3	0.6	0.3	0.2		
R11			0.5	0.3	0.3	0.4	0.4	0.4	0.1		

Note. Critical cells are in boldface. R and C refer to rows and columns, respectively, that comprise the matrix that defines the placements of the located object. Dashes indicate the location of the reference object.

Experiment 4

In Experiment 3, acceptability ratings peaked for the centermost point within the direct region and dropped off with distance from that center point. The various models can all account for such a peak, but they attribute it to different factors. According to the PC, PC-BB, and AVS models, acceptability will peak above or near the center of mass, because all three models predict a center-of-mass effect. In contrast, according to the (already disconfirmed) BB model, acceptability will peak at the midpoint: the point equidistant from the two edges of the rectangle. In Experiment 3, midpoint and center of mass were at the same location. The goal of Experiment 4 was to dissociate these points, to provide another test of the BB model against the other models. To this end, acceptability ratings were collected for four placements around a wide triangular landmark, as indicated in Figure 7. This landmark could either be upright or inverted. For both triangles, the critical points are A, B, and C. Point A is above the center of mass of the triangle, Point B is above the midpoint of the base of the triangle, and Point C is placed so that its distance from B equals the distance between A and B. The BB model predicts no difference between Points A and C, because both are equidistant from B and therefore equidistant from the edges. However, the PC, PC-BB, and AVS models all predict that Point A (center-of-mass point) will be rated significantly higher than Point C. This is because these models are sensitive to the center of mass rather than the midpoint of the landmark object.

Method

Participants. Thirty University of Notre Dame undergraduates participated in exchange for either extra credit in an undergraduate psychology class or a payment of $6 per hour of participation. None had participated in Experiments 1–3.

Stimuli. A wide triangle was created by cutting a diagonal (top left corner to bottom right) through wide landmarks from Experiment 3. It was presented both upright and inverted, as shown in Figure 7. The same four locations were probed around both triangles.

Procedure. The procedure from Experiments 1–3 was used. In all, participants performed eight trials (two landmarks × four placements).

Results and Discussion

Empirical data. Mean acceptability ratings as a function of landmark (upright vs. inverted) and trajector placement are shown in Figure 7. The critical points (A, B, and C) are in boldface. Acceptability ratings for these critical points were submitted to a 2 (landmark: upright vs. inverted) × 3 (points: A, B, C) repeated measures ANOVA. There was no effect of landmark ($F < 1$), a main effect of points, $F(2, 58) = 4.5$, $p < .015$, $MSE = 0.84$, and a significant interaction between landmark and points, $F(2, 58) = 3.9$, $p < .026$, $MSE = 0.34$. To follow up the interaction, a 95% confidence interval was constructed; a critical difference of .30 was required for significance. Thus, for the upright triangle, Point C ($M = 7.4$) was significantly lower than Points A ($M = 8.1$) and B ($M = 8.0$), which did not differ. For the inverted triangle, Point C ($M = 7.8$) was significantly lower than Point B ($M = 8.1$) but not Point A ($M = 8.0$), and Points A and B did not differ. Note, however, that the means show the same pattern as for the upright triangle. These results challenge the prediction made by the BB model that ratings will not differ between points equidistant from the midpoint; rather they support the importance of center of mass as a feature for defining spatial relations.

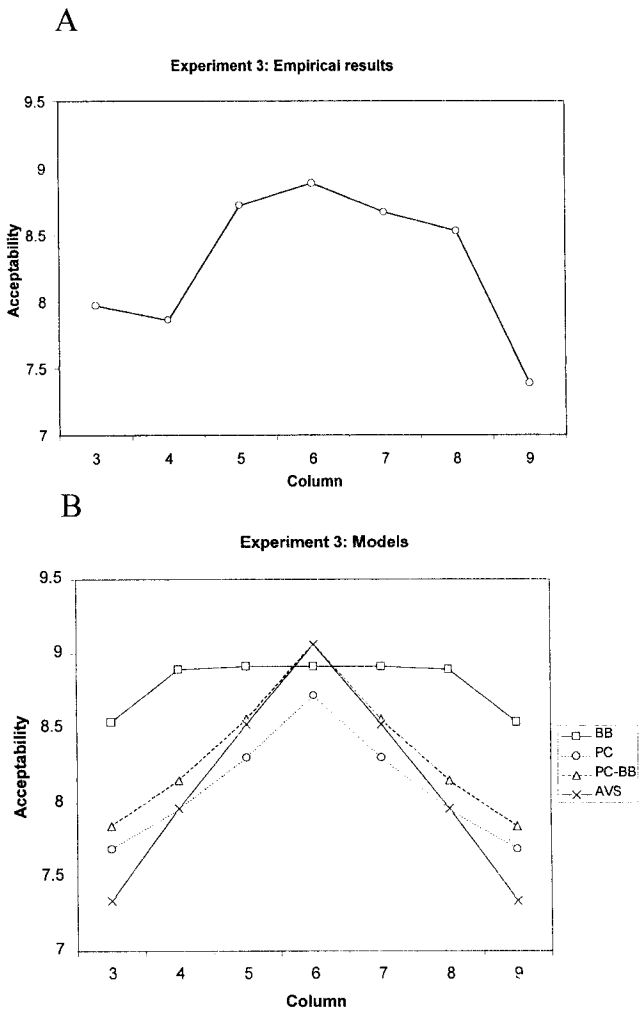

Figure 6. (A) Empirical data and (B) model results for Experiment 3. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum.

Model fits and simulations. The results of the models' fits to the data are shown in Table 5. To determine which models would exhibit the asymmetry about the midpoint obtained in the empirical data, we examined the models' predictions for Points A, B, and C relative to the upright triangle. (These predictions were qualitatively the same for the inverted triangle.) Figure 8 displays (A) the empirically determined acceptability ratings for these points and (B) the predictions of the various models.

As discussed above, the empirical data show that Point A is rated considerably higher than Point C. The PC, PC-BB, and AVS models all capture this difference. This is to be expected, because all three models are sensitive to the center-of-mass orientation, and Point A is located directly above the center of mass. The BB model, however, does not capture this aspect of the empirical data. It produces ratings that appear to be entirely flat. However, close examination of the BB-predicted ratings reveals that they are very slightly symmetrically peaked about the midpoint, so that Points A and C receive the same rating (A = 8.982, B = 8.995, C = 8.982). This midpoint peak is to be expected, because the BB model is sensitive to the midpoint rather than the center of mass. The BB model's failure to exhibit a center-of-mass effect here serves as a further disconfirmation of that model.

Summary of Experiments 1–4

Taken as a whole, these first four experiments and the accompanying model simulations provide evidence of the proximal and center-of-mass orientations in spatial term ratings. They also argue against models based on horizontal and vertical distance from the edges of the landmark. Thus, this first set of experiments supports the PC, PC-BB, and AVS models over the BB model.

Testing for the Effect of the Grazing Line

In Experiments 5–6, we attempt to further discriminate among the remaining models: PC, PC-BB, and AVS. We do this by testing a prediction of the PC-BB and AVS models: Points above the horizontal grazing line will receive higher ratings than points below it. This prediction is shared with the already-disconfirmed BB model. But critically, the PC model, unlike these others, has no explicit grazing-line mechanism and therefore might not be able to show such an effect. Thus, testing for the effect might further narrow the range of viable models.

Experiment 5

Above acceptability judgments for 65 trajector placements around an upright L-shaped landmark were collected, as shown in

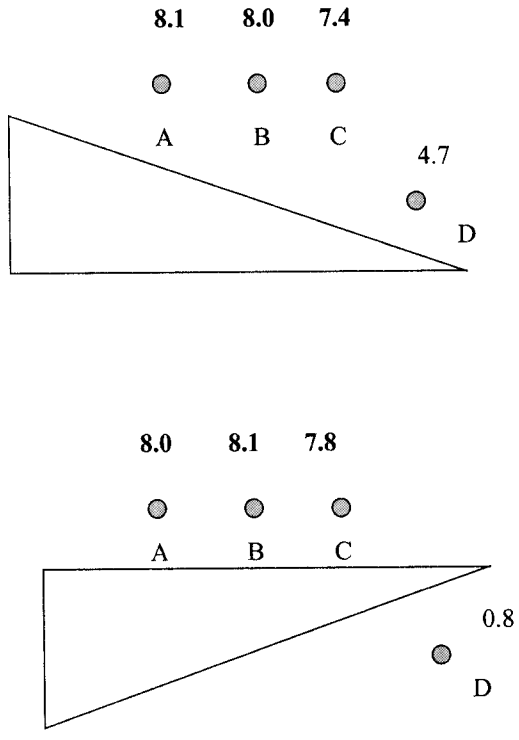

Figure 7. The upright and inverted triangles used as landmarks in Experiment 4. Trajector Point A corresponds to the center of mass of the landmark; Point B corresponds to the midpoint; Point C is the same distance from B as A is from B; Point D is within bounding box of the landmark but below the grazing line.

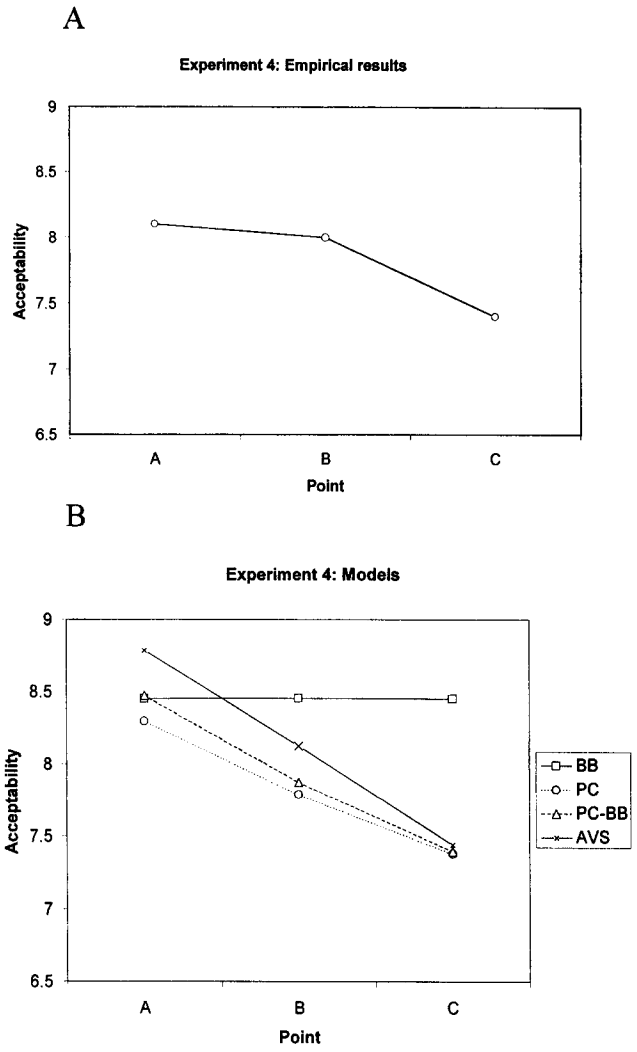

Figure 8. (A) Empirical data and (B) model results for Experiment 4. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum.

Figure 9. Of these 65 placements, we were particularly interested in comparing the ratings of points that straddled the grazing line. Thus, we compared the 6 points that were above the base of the landmark and above the grazing line with the 6 points that were above the base of the landmark but the same distance below the grazing line. These points are shown in boldface in Figure 9. If the grazing line played a role in defining spatial relations, then points above the grazing line would be rated significantly higher than points below the grazing line, with distance from the line equated.

Method

Participants. Twenty-four University of Notre Dame undergraduates participated in exchange for either extra credit in an undergraduate psychology class or a payment of $6 per hour of participation. None had participated in Experiments 1–4.

Stimuli. Upright and inverted L-shaped stimuli were used as landmarks. The height of the base and the width of the arm were both 25 mm, and the height of the arm and the length of the base were 63 mm. The

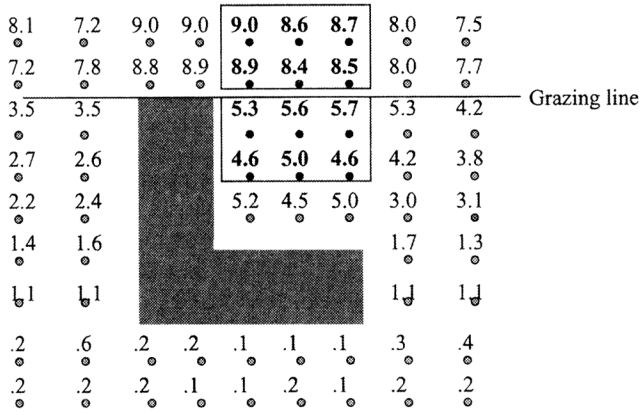

Figure 9. Placements and mean acceptability ratings around L-shaped landmark in Experiment 5. Grazing line is indicated. Critical placements are outlined and appear in boldface.

trajector from Experiment 3 was used; it was placed in the same 65 locations around each landmark, as shown in Figure 9. Only the relation *above* was tested.

Procedure. The procedure from Experiments 1–4 was used. In all, participants performed 130 trials (2 landmarks × 65 placements).

Results and Discussion

Empirical data. The inverted L-shaped stimulus did not have placements of the located object that were diagnostic of a grazing-line effect. As such, these trials were considered fillers. Mean acceptability ratings for placements of the trajector around the upright L-shape are shown in Figure 9, with critical points framed. There was a significant difference between the points above the grazing line (M = 8.7 across the six points) and the points below the grazing line (M = 5.1), $t(23)$ = 8.8, $p < .001$. These results suggest that the grazing line is an important feature for defining spatial relations.

Model fits and simulations. We presented the models with trajector placements relative to the L-shaped landmark. The results of the fits are shown in Table 5. We then focused in particular on the critical placements above and below the grazing line. For each model, we calculated the average predicted rating above the grazing line minus the average predicted rating below the grazing line. A positive resulting quantity indicated an effect of the grazing line. The results are shown in Table 8.

Table 8
Effects of the Grazing Line Relative to an L-Shaped Landmark, Following Experiment 5

Model				Empirical data
BB	PC	PC-BB	AVS	
3.050	1.652	3.640	3.738	3.528

Note. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum. Values shown are the average model-predicted ratings for selected points higher than the grazing line minus the average for points below the grazing line. A positive value indicates an effect of the grazing line. All models exhibit the effect.

All four models exhibit an effect of the grazing line. For the PC-BB and AVS (and the disconfirmed BB) models, this is to be expected, because the grazing line is an integral part of the architecture of these models. The PC model, in contrast, has no such element in its design, but it too exhibits the effect. This must be attributed to the fact that the proximal and center-of-mass orientations also varied across the critical points of the experiment. Thus, this simulation fails to qualitatively discriminate among the models. Quantitatively however, the overall measure of fit suggests that the PC model does not account as well for these data as the PC-BB and AVS models.

Experiment 6

Experiment 5 provided initial evidence of the importance of the grazing line. However, in that experiment, center-of-mass and proximal orientations were varied across the critical points. Because these features were not controlled, the effect observed in Experiment 5 might be due to these features and not to placement relative to the grazing line. To address this concern, in Experiment 6, *above* acceptability judgments were collected for 31 trajector placements around a tall triangular landmark, as shown in Figure 10. Of these 31 locations, there were two critical pairs of points (framed in Figure 10) that satisfied the following criteria. First, the center-of-mass and proximal orientations of points within a pair were equal. Second, one member of the pair was located above the grazing line, and the other was located below the grazing line.[9] Given these criteria, any differences between points within a line could be attributed to an influence of the grazing line, but not to the center-of-mass or proximal orientations.

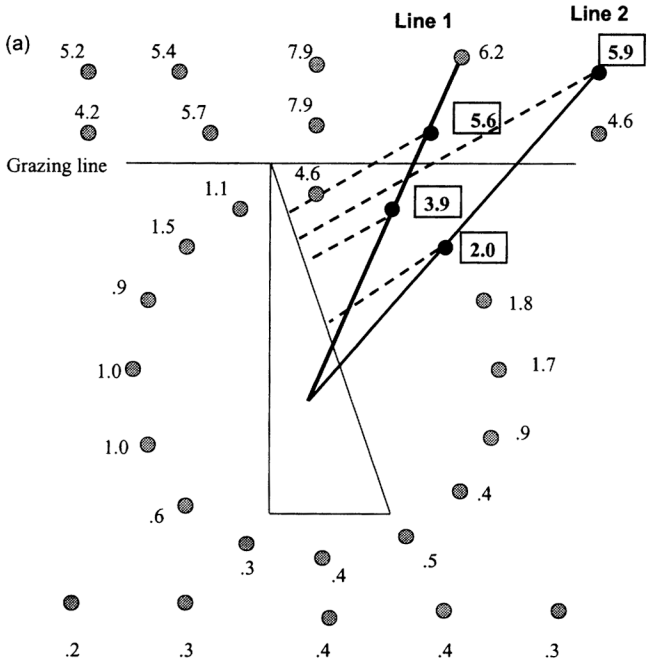

Figure 10. Placements and mean acceptability ratings around the tall triangle landmark used in Experiment 6. The four critical points on the two radiating lines are outlined and appear in boldface. Solid radiating lines indicate center-of-mass orientation. Dashed lines indicate proximal orientation.

Method

Participants. Twenty-six University of Notre Dame undergraduates participated in exchange for either extra credit in an undergraduate psychology class or a payment of $6 per hour of participation. None had participated in Experiments 1–3.

Stimuli. Tall and wide triangles were used as landmarks; these were created by cutting a diagonal (top left corner to bottom right) through the tall and wide landmarks from Experiment 3. The trajector from Experiment 3 was used; it was placed in the same 31 locations around each landmark. These are shown surrounding the tall landmark in Figure 10. Only the relation *above* was tested.

Procedure. The procedure from Experiments 1–3 was used. In all, participants performed 62 trials (2 landmarks × 31 placements).

Results and Discussion

Empirical data. The wide triangle landmark did not have placements of the located object that were diagnostic of a grazing-line effect. As such, these trials were considered fillers. Mean acceptability ratings for placements of the trajector around the tall triangle are shown in Figure 10. The critical points are framed. To determine whether there was a grazing-line effect, we submitted the acceptability ratings for these 4 cells to a 2 (grazing line: above vs. below) × 2 (line: 1 vs. 2) repeated measures ANOVA. There was a main effect of grazing line, $F(1, 25) = 25.8$, $p < .001$, $MSE = 7.94$, a main effect of line, $F(1, 25) = 4.3$, $p < .049$, $MSE = 3.61$, and a significant interaction, $F(1, 25) = 19.2$, $p < .001$, $MSE = 1.57$. To follow up the interaction, we constructed a 95% confidence interval on the basis of the error term for the interaction; a critical difference of .70 was required for significance. Thus, ratings were significantly higher when the point was above the grazing line than when the point was below the grazing line, both for the points on Line 1 (*M*s = 5.6 vs. 3.9, respectively) and for those on Line 2 (*M*s = 5.9 and 2.0, respectively). This effect was present despite the fact that the pairs of points along these lines shared the same center-of-mass and proximal orientations. Thus, the sufficiency of these orientations to predict spatial relational use is questionable; rather, a grazing-line feature might also be necessary.

Model fits and simulations. The results of the models' fits to the data are shown in Table 5. We tested the models using the stimuli of Experiment 6, focusing in particular on the two critical radiating lines, collecting model-predicted ratings at both points along each line—one point above the grazing line and one below. The value of interest was the rating for the higher point minus the rating for the lower point, with a positive value indicating a grazing-line effect. The results are displayed in Table 9.

The PC model failed to exhibit a grazing-line effect. This was to be expected. The PC model relied only on the center-of-mass and proximal orientations, and these features were held constant along the radiating lines of the experiment. Thus, we expected no difference in predicted rating between points along a radiating line: The difference scores would be zero. The difference scores actu-

[9] These displays were inspired by observations made by Annette Herskovits.

Table 9
Grazing-Line Effect Following Experiment 6

Location	Model				Empirical data
	BB	PC	PC-BB	AVS	
Line 1	1.975	−0.003	2.541	3.037	1.730
Line 2	−0.078	0.048	2.987	3.265	3.885

Note. BB = bounding box; PC = proximal and center of mass; PC-BB = hybrid; AVS = attentional vector sum. Each value shown is the model-predicted rating for the higher point along a radiating line, minus the model-predicted rating for the lower point along the same line. A positive number indicates a grazing-line effect.

ally obtained from the model were nearly, but not exactly, zero. The minor deviation resulted from the model's internal representation of the triangle as a set of points in the two-dimensional plane. This representation required that the continuous hypotenuse be discretized into a finite set of points. This discretization gave rise to minor differences across points in the calculation of the proximal orientation, hence, the minor difference in ratings. The critical point was that the PC model failed to show an effect of grazing line when orientational features were held constant. This disconfirmed the PC model. Moreover, it disconfirmed any implementation of the model, as long as it was based solely on the center-of-mass and proximal orientations.

The PC-BB and AVS models both showed a clear effect of the grazing line. Again, this was to be expected. Both models explicitly relied on two components: (a) a height function borrowed from the BB model that highlighted the grazing line and (b) an orientational element. The proximal and center-of-mass orientations were controlled for in this experiment, so that the influence of the grazing line was isolated. This influence was captured in the simulations.

(The already-disconfirmed BB model showed the effect for Line 1, but not for Line 2. This failure might seem surprising, because the BB model was by definition sensitive to the grazing line. However, it was also sensitive to horizontal distance. In Line 2, the higher point was horizontally quite far from the landmark bounding box. Thus, this point received a very low rating, even lower than the rating for the other point on the line, which was below the grazing line but near the bounding box. Because the higher point received a lower rating, the model failed to exhibit the effect. We did not take this as a further strong disconfirmation of the model, however. The BB model would not encounter this problem, given other parameter settings. Thus, we viewed this result as a minor implementational problem rather than a clear falsification. However, Experiments 2 and 4 had already disconfirmed this model in its generality, not merely in this implementation.)

Testing for the Effect of Distance: Experiment 7

Across six experiments, only two models have passed all qualitative empirical tests and have also provided tight quantitative fits to the empirical data: AVS and PC-BB. The goal of Experiment 7 was to discriminate between these two remaining models by focusing on the contrasting predictions that the models make concerning the effect of distance on acceptability ratings. The AVS model predicts that if one were to move a trajector back and forth over the surface of an object, but at a low elevation, ratings would not be very sensitive to the centeredness of the trajector above the landmark. However, at a higher elevation, there would be greater sensitivity to centeredness. The PC-BB model makes exactly the opposite prediction: Ratings in the lower row would be more sensitive to centeredness. Thus, these two models can be distinguished by comparing ratings for placements of the trajector that are very close to the landmark with ratings for placements that are far from the landmark.

Method

Participants. Forty-seven University of Notre Dame undergraduates participated in exchange for either extra credit in an undergraduate psychology class or a payment of $6 per hour of participation. None had participated in Experiments 1–6.

Stimuli and procedure. The wide rectangle landmark from Experiment 3 was used. The trajector was placed at 25 locations surrounding the landmark (see positions indicated by the acceptability ratings in Table 10). Only the relation *above* was tested. The procedure from Experiments 1–4 was used.

Results and Discussion

Empirical data. Mean acceptability ratings for *above* for placements of the trajector around the wide landmark are given in Table 10. The cells of interest are in rows 1–2, columns 2–8. Mean acceptability ratings for these cells were regressed on the absolute value of the center-of-mass orientations and distance (explicitly coded as 1 or 2, corresponding to rows 1–2) for the 14 cells using linear regression. The regression equation revealed a significant influence of center of mass (beta weight = −4.09, $p < .002$) and a significant interaction between center of mass and row (beta weight = 3.87, $p < .008$). There was no effect of distance ($p > .15$, overall model $R^2 = .76$). As shown in Table 10, the effect of center of mass was greater for trajector placements that were far from the landmark (row 1) than for trajector placements that were close to the landmark (row 2), as predicted by AVS.

Model fits and simulations. The model fits for this experiment are shown in Table 5. under "Experiment 7, Wide rectangle." To qualitatively test the models for a distance effect, we recorded each model's predicted values for the critical stimuli used in the experiment. As in the experiment, we focused on points in the direct region, in columns 2–8, rows 1 and 2. We looked for centeredness effects within these two rows. The results are displayed in Figure 11. The

Table 10
Mean Ratings for Above by Placement for Experiment 7

Row	Column								
	C1	C2	C3	C4	C5	C6	C7	C8	C9
R1	6.3	**7.8**	**8.3**	**8.5**	**8.8**	**8.4**	**8.2**	**7.7**	6.2
R2	4.7	**7.7**	**8.1**	**8.3**	**8.3**	**8.3**	**8.3**	**7.7**	4.3
R3	1.3	—	—	—	—	—	—	—	1.4
R4	0.8		0.9		0.7		0.7		0.8

Note. Critical cells are in boldface. The spacing of the rows reflects the distance manipulation. R and C refer to rows and columns, respectively, that comprise the matrix that defines the placement of the located object. Dashes indicate the location of the reference object.

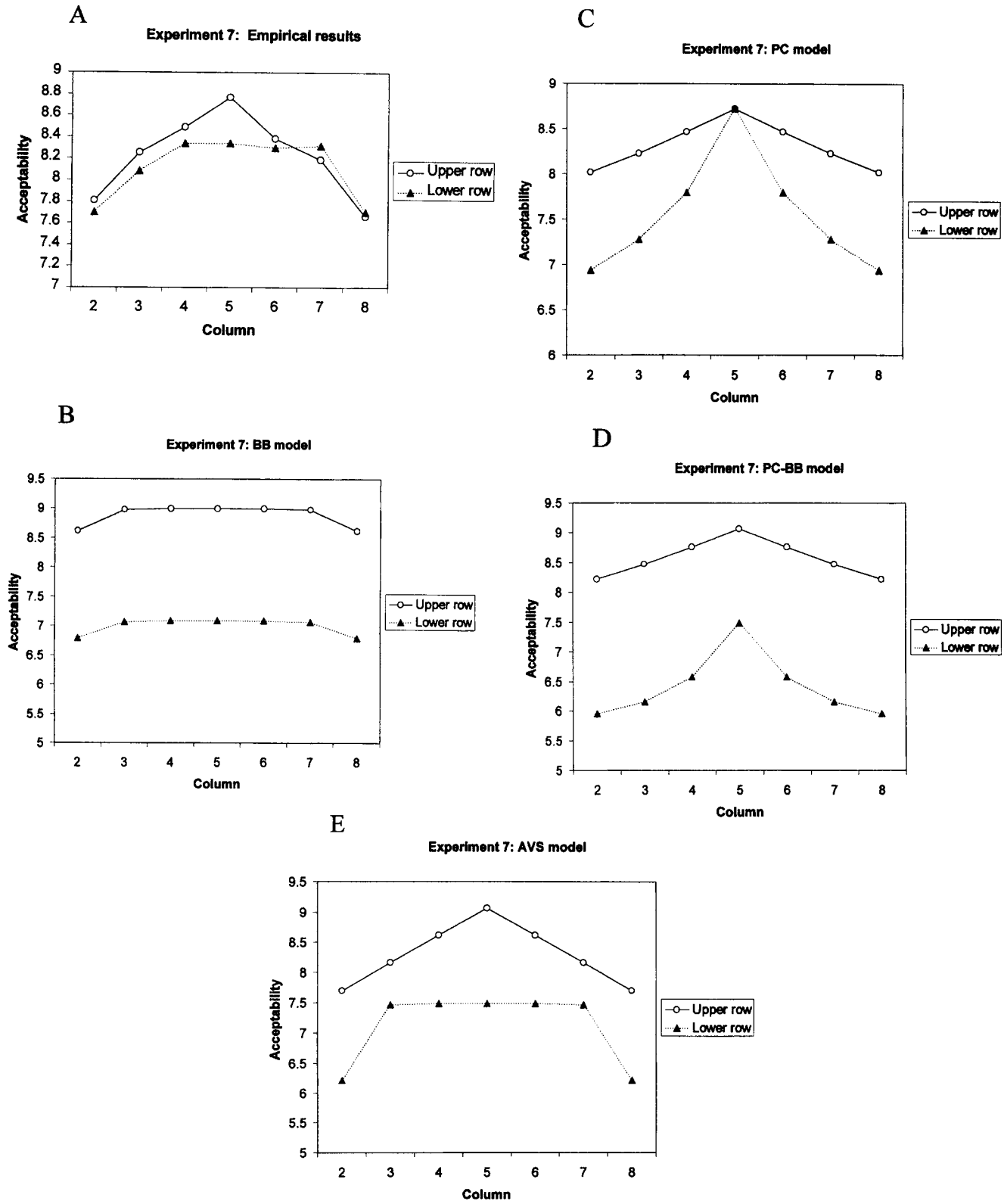

Figure 11. Results of Experiment 7. Panel A shows empirical data; Panels B–E show predicted output from bounding box, proximal and center-of-mass, hybrid, and attentional vector-sum models, respectively.

empirical data are displayed in A, and the predictions of the different models are shown in subsequent frames. The empirical data show a clear peak in the ratings of the upper row, replicating Experiment 3. But the ratings in the lower row are relatively flat. Thus, as the AVS model leads us to expect, there is a greater effect of centeredness for high points than for low points. The PC-BB model fails to exhibit this effect. Instead, it predicts flatter ratings for the higher row than for the lower row, the opposite of what we find experimentally. The already-disconfirmed BB and PC models also fail to exhibit the effect. The PC model behaves as the PC-BB model does, and the BB model shows largely flat responses at both elevations.

The AVS model does exhibit the predicted qualitative effect: The ratings are more peaked in the upper row than in the lower row. And as shown in Table 5, its fit to the data is better than that of the other models. This is true when all 25 data points are considered ("Experiment 7, wide rectangle") and also when we restrict our attention to the 14 critical points shown in the figure. This supports the AVS model over its competitors.

However, the AVS model does not capture the details of these data perfectly. In the AVS simulation, the ratings for the upper row are all considerably higher than the ratings for the corresponding points in the lower row. In the empirical data, in contrast, this is not always the case. This raises a question: Is the AVS model capable of providing a closer fit to the Experiment 7 data, under other parameter settings? If not, its status as the survivor of our competition could potentially be called into question. To test this, we adjusted parameters to optimize the fit of each model directly to the Experiment 7 data. The resulting parameter values are shown in Table 1 and the fits are shown in Table 5. As before, numbers are given for the fit to the Experiment 7 data set as a whole ("Fit directly to data") and for the critical points only. Again, the AVS model provides the best overall fit. The BB fit is very nearly as good, but that model has already been disconfirmed on other grounds. Significantly, when we examine the critical points only, the AVS model clearly provides the best fit—a much tighter fit than obtained with the original parameter settings and tighter than that of any other model. Figure 12 displays each

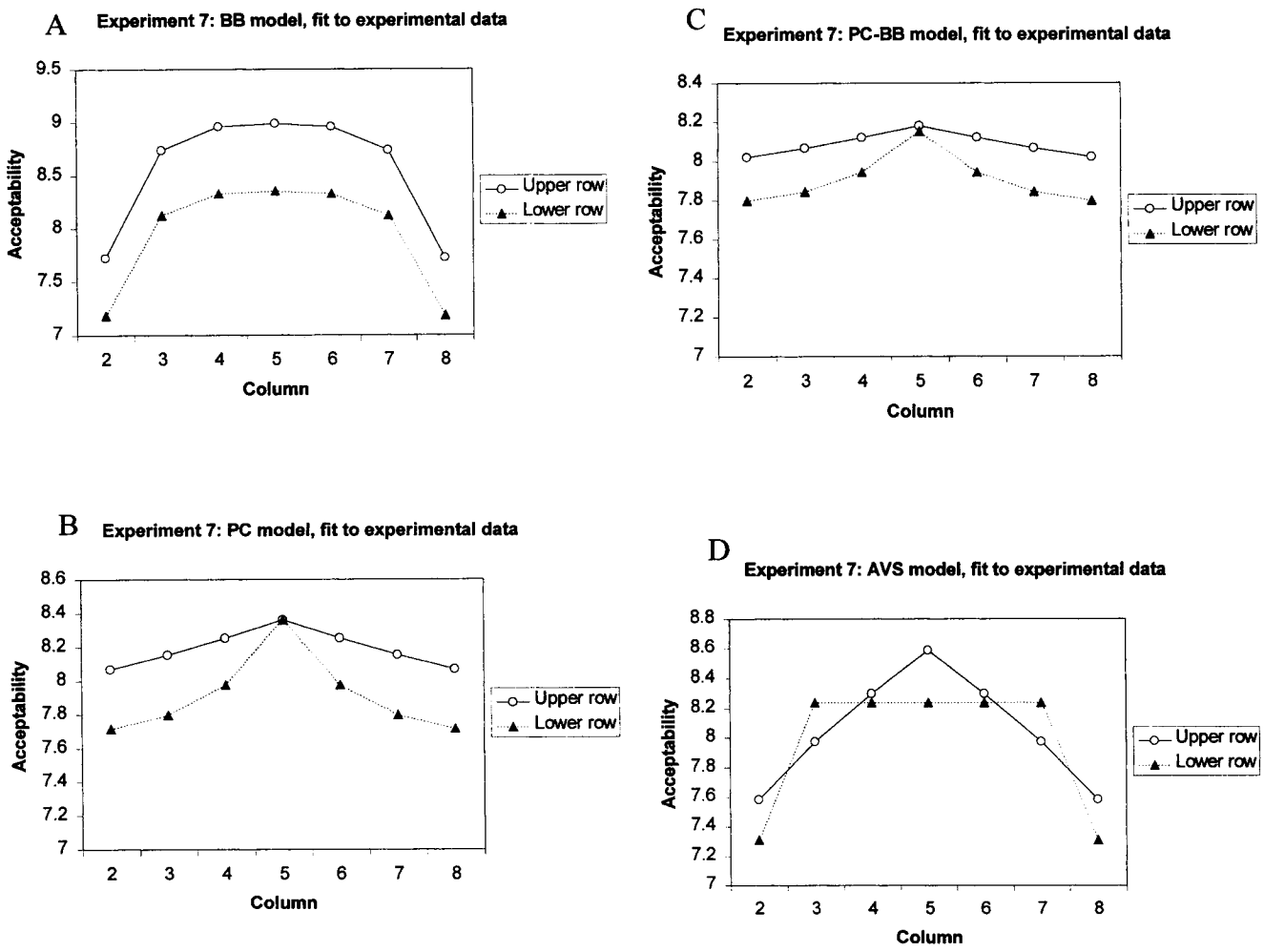

Figure 12. Results of Experiment 7, with model parameters set specifically to fit these data. Panels A–D show predicted output from bounding box, proximal and center-of-mass, hybrid, and attentional vector-sum models respectively.

model's new fit to the critical points from this experiment. As before, the AVS model exhibits the qualitative effect of peakedness as a function of distance while also providing a fairly close fit to these points overall. The other models again fail to exhibit the effect, even when fit directly to these data. Critically for our purposes, these findings support the AVS model and also disconfirm the PC-BB model, the only remaining competitor under consideration.

Composite Model Fits

As a final test of the models, we examined their fits to all of our experimental *above* data, pooled together. Here, we again used model parameter values obtained from fitting Logan and Sadler's (1996) data, not our own. This yielded a set of overall composite fits, shown in Table 5. All four models performed reasonably well, with the AVS model providing the best fit. Figure 13 shows these composite fits. All empirically obtained *above* data from Experiments 1–7, with the predictions of each model and descriptions of the stimuli, are available at http://www.ccp.uchicago.edu/~regier/avs.

Individual Differences

The AVS model also accounts for some individual differences in patterns of *above* ratings. We examined the responses of individual participants and found that a good proportion of them (22%–58%, depending on the experiment) gave a maximal response as long as the trajector was directly above some part of the landmark. Thus, these individuals gave responses that were flat across the direct region; however, other individuals did not provide such flat responses. The AVS model accounts for both response patterns, with different parameter settings. If the attentional beam-width free parameter is extremely narrow, there is only negligible integration over the landmark, regardless of height. Under these circumstances, the vector sum will be perfectly aligned with upright vertical as long as the trajector is above some part of the landmark—yielding flat ratings of the sort observed. But when the

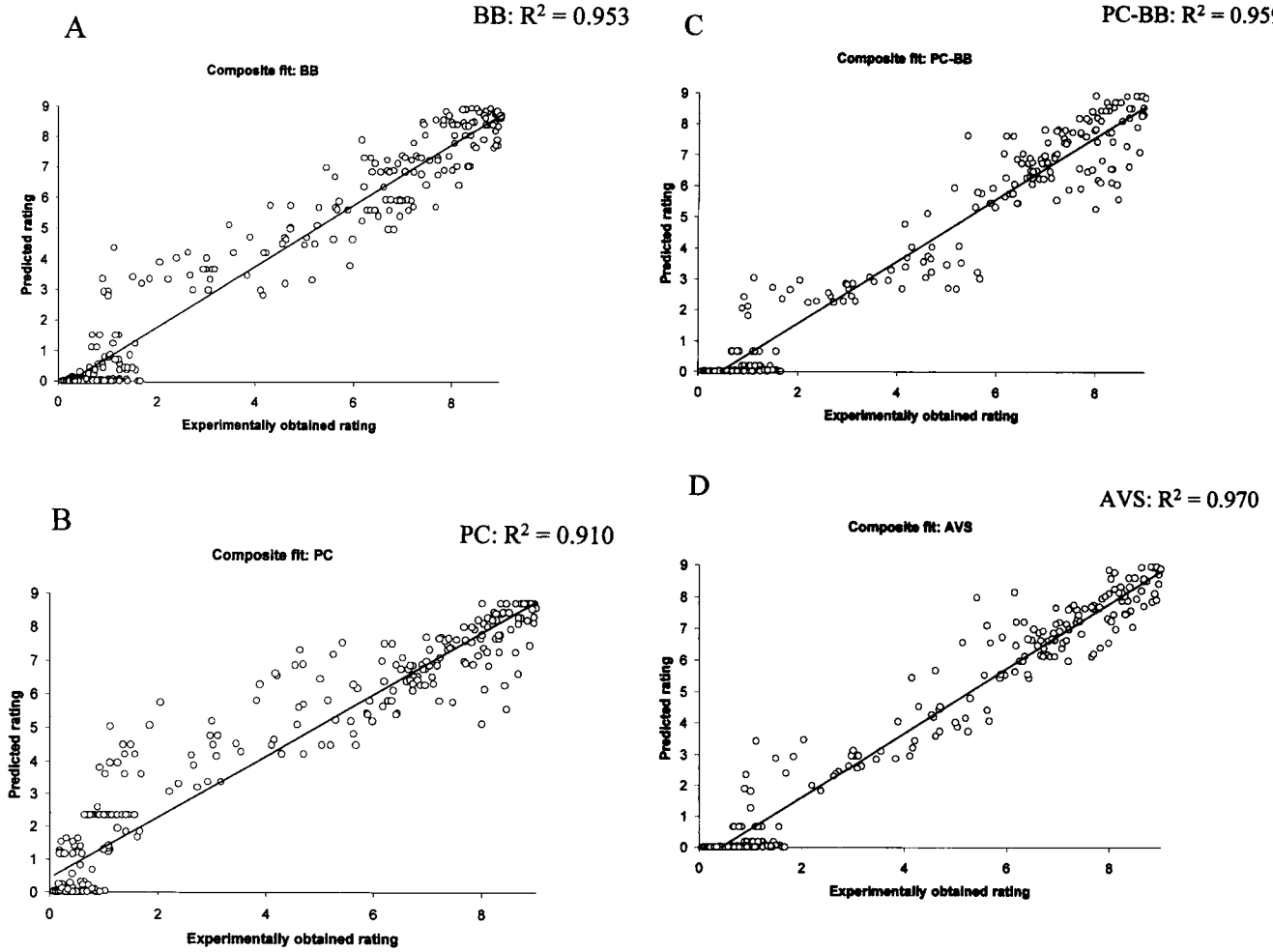

Figure 13. Composite model fits to pooled data from all experiments (1–7). (A) Bounding box model (R^2 = .953). (B) Proximal and center-of-mass model (R^2 = .910). (C) Hybrid model (R^2 = .959). (D) Attentional vector-sum model (R^2 = .970).

attentional beam is broader, the AVS model integrates over the landmark, predicting nonflat ratings of the sort observed. With different parameter settings, the AVS model exhibits the qualitative effects shown in both the flat and nonflat responses and provides good quantitative fits to both response types.

No model other than AVS accounts for both response patterns. However, the BB and PC-BB models can account for the flat responses, with appropriate settings of their parameters. Thus, despite their elimination in our model competition, there is partial support for these two models: They account for a subclass of individuals. The PC model, in contrast, cannot consistently account for either class of response. Details may be found at http://www.ccp.uchicago.edu/~regier/avs/diffs-results.html. Thus, both aggregate responses and individual differences in response provide support for the AVS model.[10]

General Discussion

Our studies confirm the three central predictions of the AVS model. First, spatial term ratings are influenced by the proximal and center-of-mass orientations. Second, ratings are sensitive to the grazing line. Third, ratings are affected by distance. The AVS model also provides a reasonably good fit to our experimental data. In contrast, none of its competitors pass all empirical tests, and none fit the data as well.

What is the significance of this finding? What does it tell us about the relation between language and perception? The AVS model is motivated in part by the role of attention in spatial perception and by vector-sum coding of direction. Thus, to the extent that its success stems from these independently motivated sources, it provides a preliminary grounding of linguistic spatial categories in nonlinguistic perception.

As it happens, its success stems only in part from these sources. Recall that the AVS model consists of two components: the attentional vector sum, which is independently motivated, and the height function, which is not. This height function was borrowed from the BB model for purely data driven reasons. It explicitly lowers ratings for points below the grazing line, yielding a grazing-line effect. Thus, we cannot attribute the AVS grazing-line effect to independently motivated aspects of the model.

However, the effects of proximal and center-of-mass orientation and the distance effect are attributable to the attentional vector sum. As we have seen, these effects emerge from the interaction of attention and vector-sum coding. In addition, much of the model's success in quantitative fits to the data may also be attributed to the attentional vector sum. Specifically, the attentional vector sum accounts for most of the variation in ratings for trajector positions that are well above the grazing line. We know this because the height function is almost uniform across these positions. The critical point is that significant aspects of the AVS model's success may be traced to its independently motivated components. Thus, it grounds some aspects of spatial language in perception.

We are not the first to put forth such a proposal. Landau and Jackendoff (1993) noted that spatial terms tend to encode the nature of the related objects in a sketchy or schematic fashion, specifying only their gross perceptual characteristics. They proposed that this fact reflects the underlying neural representation of spatial relations. In particular, they noted that the visual system is divided into two cortical pathways, one capturing where an object is and the other capturing what it is (Ungerleider & Mishkin, 1982). They suggested that the schematic nature of object representation in spatial term semantics might result from this neural division of labor. Because spatial terms code location, they are presumably underpinned by the *where* pathway, and this pathway codes the objects themselves in less detail than does the *what* pathway.

Our argument bears similarities to this one and differences from it. The most important similarity is that our overall point is essentially the same: Linguistic spatial categories can be explained in terms of underlying structures that are not linguistic in character. We feel there is an important difference, however, concerning level of detail. A strength of the AVS account is that it is quite fine-grained, relying on quantitative matches to empirical data rather than perceived conceptual commonality between an aspect of spatial language and an aspect of visual system architecture. It is this fine grain that allowed us to discriminate between the AVS account and other plausible-seeming accounts against which it was pitted. Of course, we readily acknowledge that we sacrifice the breadth of coverage that Landau and Jackendoff (1993) have achieved.

The work of Hayward and Tarr (1995) is another natural comparison. They collected spatial term acceptability judgments in the same manner that we did and noted participants' performance on a related but nonlinguistic spatial memory task. They found that the best performance on the spatial memory task occurred at the prototypes, or best examples, of the linguistic categories (i.e., in the direct region above, below, to the left, or to the right of the landmark). They concluded that because performance peaked at the same locations for both their linguistic and nonlinguistic tasks, the same structures might subserve both tasks. Crawford, Regier, and Huttenlocher (2000) further examined this issue, using similar tasks and additional analyses. Although their findings argue against some of Hayward and Tarr's (1995) claims, they further supported the general notion that structure is shared across linguistic and nonlinguistic spatial categories. Again, the overall point is a similar one: the possible grounding of spatial language in spatial perception.

It is also natural to compare this result with earlier work in the semantic domain of color. Berlin and Kay (1969) found that different languages have different color categories. For example, not every language has a word directly translatable as English *purple*. More interestingly, they also found substantive constraints on this crosslinguistic variation. They found that the best examples of color terms, across languages, occurred at the same 11 foci along the color spectrum, regardless of the color categorization

[10] There is a possible qualification to our claim that the AVS model accounts well for our data, both individual and aggregate. Like the other models, AVS predicts left–right symmetry of ratings relative to symmetrical landmarks. It is not clear whether this prediction is entirely valid. In Experiments 2 and 3, points on the right side of space received very slightly higher ratings than corresponding points on the left. The mean left–right difference was -0.1236, $t(71) = -3.2786$, $p = .0016$. We found no further evidence of such an asymmetry, however. In fact Experiment 7 showed a nonsignificant trend in the opposite direction ($M = 0.08$), $t(10) = 1.70$, $p = .12$. Given the inconsistent nature of these findings and the extremely small size of the asymmetry we did find, we are comfortable advancing the symmetry-predicting AVS model.

scheme of the language in question. Kay and McDaniel (1978) presented a reduction of these linguistic color foci to the neurophysiology of the visual system. In particular, they noted a physiologically observable neural coding for the 4 color categories red, yellow, green, and blue in the opponent response cells of the lateral geniculate nucleus and suggested a straightforward mechanism for building the other color foci up out of these. Thus, their work suggested that the observed constraints on crosslinguistic variation result from the fact that all humans share essentially the same perceptual apparatus.

Although we have not yet explored the crosslinguistic applicability of the AVS model, we are eager to do so. Languages vary in their structuring of space, but there are limits to this variation (Bowerman, 1996; Talmy, 1983), just as there are in the domain of color. Because the AVS model is grounded in putatively universal perceptual processes, it would bolster our argument if the spatial term acceptability judgments of speakers of other languages also appeared to implicate such processes. For instance, the AVS model appears to be consistent with German data collected by Gapp (1995) but with different parameter settings from those dictated by our English data. We intend to investigate this further in the near future.

Limitations and Extensions

The work as it stands is limited in several important respects and calls for extension.

Trajector Shape

The shape of the trajector is limited to a single point in the current model. A simple extension to this model would allow for extended trajectors. Such a new model would use vectors rooted at the landmark, pointing to the closest point on the trajector. This is clearly equivalent to the current model in the point-trajector case and may be adequate for other trajector shapes as well. We intend to investigate this in future research.

Object Function

It has been shown that the functional relationship between the landmark and trajector affects the manner in which spatial configurations are described (Carlson-Radvansky, Lattanzi, & Covey, 1999; Carlson-Radvansky & Tang, 2000; Carlson-Radvansky & Radvansky, 1996; Coventry, Carmichael, & Garrod, 1994; Coventry & Prat-Sala, 1998). For example, Carlson-Radvansky et al. elicited *above* ratings for a coin relative to a piggy bank. They found that people gave the highest ratings when the coin was placed above a functionally important part of the piggy bank: the slot through which it was meant to drop. This was found even when this functionally important part was dissociated from the center of mass of the piggy bank. More specifically, across participants, the slot of the piggy bank was moved, so that sometimes it was at the back of the pig, sometimes it was in the middle, coinciding with its center of mass, and sometimes it was at the front of the pig. The critical finding was that the peak in the acceptability ratings for placements of the coin above the piggy bank shifted in accordance with the position of the slot.

Functional features are notably absent in our displays, and thus the AVS model does not currently address function. We consider this a serious limitation. However, we also feel that a natural extension to the model might be able to account for functional effects. Specifically, the functional part of an object might receive greater amounts of attention than the remainder of the object. There is some support for this notion: Lin and Murphy (1997) have shown that people more quickly and accurately detect that a part is missing from an object when it is a functional part as opposed to a nonfunctional part. Thus, the function of an object part would bias the attentional vector sum, which would return higher ratings for positions near the functional part of the object. In this manner, some functional effects may be explicable in terms of the AVS model. A formal version of this extension of the model is currently being explored.

Lexical Competition

Finally, another issue concerns the height function used by the AVS model. As we have seen, this part of the model is somewhat ad hoc, in that it was chosen to directly reflect the data rather than explain it. Thus, it is the "unmotivated" component of the model. One possibly relevant observation in this regard concerns *lexical competition* (MacWhinney, 1987). This is the notion that the applicability of one term in the lexicon is affected by the applicability of other related terms. For example, a trajector that is below a landmark provides a poor example of *above* precisely because it is *below:* The two terms are opposites. Humans do have to select among contending lexemes, and there is currently no locus for such a competition in the AVS model. We speculate that at least some aspects of the height function we are currently using may be explicable in terms of lexical competition between opposing terms. This is another issue for future studies to address.

Effects and Mechanisms

One general, broadly applicable conclusion that may be drawn from this work is distinctly cautionary in character. This conclusion is that effects need not reflect underlying mechanisms in a straightforward, transparent fashion. This point has been made several times recently, particularly in the connectionist literature (Elman et al., 1996; Rumelhart & McClelland, 1986; Seidenberg & McClelland, 1989). We add our voices to this chorus. In our case, the PC and PC-BB models account for the effects of proximal and center-of-mass orientations in a straightforward fashion: The effects are essentially built into the architecture. Ultimately, however, we have concluded that this easy answer is the wrong one. One reason for this decision is that the easy models are not tight; they do not fit our data as well as does the competing AVS model. Another reason is that the easy model is not particularly illuminating. To enshrine effects as mechanisms would be to make do with a descriptive model of the phenomenon, without any explanatory links to other known phenomena. The AVS model, in contrast, does provide such links while accounting for the same empirical effects and others. In so doing, this latter model gives us an initial grounding of some aspects of spatial language in the nonlinguistic structures and processes of visual perception.

References

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Brown, P., & Levinson, S. C. (1993). Uphill and downhill in Tzeltal. *Journal of Linguistic Anthropology, 3,* 46–74.

Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition, 46,* 223–244.

Carlson-Radvansky, L. A., & Irwin, D. E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language, 33,* 646–671.

Carlson-Radvansky, L. A., Lattanzi, K., & Covey, E. (1999). "What" effects on "where": Functional influences on spatial relations. *Psychological Science, 10,* 516–521.

Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language, 37,* 411–437.

Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science, 7,* 56–60.

Carlson-Radvansky, L. A., & Tang, Z. (2000). Functional influences on orienting a reference frame. *Memory & Cognition, 28,* 812–820.

Casad, E. H. (1988). Conventionalization of Cora locationals. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (Vol. 50, pp. 345–378). Philadelphia: Benjamins.

Clark, H. H. (1973). Space, time, semantics, and the child. In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). San Diego, CA: Academic Press.

Coventry, K., Carmichael, R., & Garrod, S. C. (1994). Spatial prepositions, object-specific function, and task requirements. *Journal of Semantics, 11,* 289–309.

Coventry, K., & Prat-Sala, M. (1998). Geometry, function, and the comprehension of over, under, above and below. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 261–266). Mahwah, NJ: Erlbaum.

Crawford, E., Regier, T., & Huttenlocher, J. (2000). Linguistic and nonlinguistic spatial categorization. *Cognition, 75,* 209–235.

Downing, C., & Pinker, S. (1985). The spatial structure of visual attention. In M. Posner & O. Marin (Eds.), *Attention and performance XI* (pp. 171–187). Hillsdale, NJ: Erlbaum.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Emmorey, K., & Casey, S. (1995). A comparison of spatial language in English and American Sign Language. *Sign Language Studies, 88,* 255–288.

Farah, M. J., Brunn, J. L., Wong, A. B., Wallace, M. A., & Carpenter, P. A. (1990). Frames of reference for allocating attention to space: Evidence from the neglect syndrome. *Neuropsychologia, 28,* 335–347.

Franklin, N., Henkel, L. A., & Zengas, T. (1995). Parsing surrounding space into regions. *Memory & Cognition, 23,* 397–407.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119,* 63–76.

Gapp, K.-P. (1994). A computational model of the basic meanings of graded composition spatial relations in 3-D space (Tech. Rep. III). Department of Computer Science: Universität des Saarlandes.

Gapp, K.-P. (1995). Angle, distance, shape, and their relationship to projective relations. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 112–117). Mahwah, NJ: Cognitive Science Society.

Garnham, A. (1989). A unified theory of meaning of some spatial relational terms. *Cognition, 31,* 45–60.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science, 223,* 1416–1419.

Harris, C. (1990). Connectionism and cognitive linguistics. *Connection Science, 2,* 7–34.

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition, 55,* 39–84.

Helmantel, M. (1998). Simplex adpositions and vector theory. *The Linguistic Review, 15,* 361–388.

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

Herskovits, A. (1998). Schematization. In P. Olivier & K.-P. Gapp (Eds.), *Representation and processing of spatial expressions* (pp. 149–162). Mahwah, NJ: Erlbaum.

Janda, L. (1988). The mapping of elements of cognitive space onto grammatical relations: An example from Russian verbal prefixation. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (Vol. 50, pp. 327–343). Philadelphia: Benjamins.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language, 54,* 610–646.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101–124.

Landau, B., & Jackendoff, R. (1993). What and where in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16,* 217–265.

Langacker, R. (1987). *Foundations of cognitive grammar I: Theoretical prerequisites.* Stanford, CA: Stanford University Press.

Lee, C., Rohrer, W. H., & Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature, 332,* 357–360.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van der Grind, & J. J. Koenderink (Eds.), *Limits in perception* (pp. 323–358). Utrecht, the Netherlands: VNU Science Press.

Levinson, S. (1996). Frames of reference and Molyneux's questions: Cross-linguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Lin, E. L., & Murphy, G. L. (1997). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1153–1169.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1015–1036.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28,* 103–174.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 493–529). Cambridge, MA: MIT Press.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Erlbaum.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Munro, P., Cosic, C., & Tabasko, M. (1991). A network for encoding, decoding and translating locative prepositions. *Connection Science, 3,* 225–240.

Nenov, V., & Dyer, M. (1994). Perceptually grounded language learning: II. DETE: A neural/procedural model. *Connection Science, 6,* 3–41.

Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism.* Cambridge, MA: MIT Press.

Regier, T. (1997). Constraints on the learning of spatial terms: A compu-

tational investigation. In R. Goldstone, P. Schyns, & D. Medin (Eds.), *Psychology of learning and motivation: Mechanisms of perceptual learning* (Vol. 36, pp. 171–217). San Diego, CA: Academic Press.

Rumelhart, D., & McClelland, J. (1986). On learning the past tenses of English verbs. In J. McClelland, J. D. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 216–271). Cambridge, MA: MIT Press.

Schirra, J. (1993). A contribution to reference semantics of spatial prepositions: The visualization problem and its solution in VITRA. In C. Zelinsky-Wibbelt (Ed.), *The semantics of prepositions: From mental processing to natural language processing* (pp. 471–515). Berlin: Mouton de Gruyter.

Seidenberg, M., & McClelland, J. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Shallice, T. (1996). The language-to-object perception interface: Evidence from neuropsychology. In P. Bloom, M. A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 531–552). Cambridge, MA: MIT Press.

Siskind, J. (1994). Grounding language in perception. *Artificial Intelligence Review, 8,* 371–391.

Stein, J. F. (1992). The representation of egocentric space in the posterior parietal cortex. *Behavioral and Brain Sciences, 15,* 691–700.

Suppes, P., Liang, L., & Boettner, M. (1992). Complexity issues in robotic machine learning of natural language. In L. Lam & V. Naroditsky (Eds.), *Modeling complex phenomena* (pp. 102–107). New York: Springer-Verlag.

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research and application* (pp. 225–282). New York: Plenum.

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language, 35,* 371–391.

Taylor, J. (1988). Contrasting prepositional categories: English and Italian. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (Vol. 50, pp. 299–326). Philadelphia: Benjamins.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15–48.

Ungerleider, L., & Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. Goodale, & R. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

van der Zee, E. (1996). *Spatial knowledge and spatial language.* Unpublished doctoral dissertation, Interdisciplinair Sociaal-Wetenschappelijk Onderzoeksinstituut Rijksuniversiteit/Utrecht University, Utrecht, the Netherlands.

Wilson, H. R., & Kim, J. (1994). Perceived motion in the vector sum direction. *Vision Research, 34,* 1835–1842.

Appendix A

Demonstration That the Attentional Vector-Sum Model Produces the Center-of-Mass Orientation When the Attentional Field Is of Uniform Strength

Let α be the uniform attentional strength, (x_{tr}, y_{tr}) be the location of the trajector, (x_i, y_i) be the location of point i of the landmark object, and n be the number of points in the landmark. Then the attentional vector-sum model's vector sum is as follows:

$$\sum_i a_i\vec{c}_i = \sum_i \alpha\vec{c}_i = \alpha \sum_i \vec{c}_i = \alpha \sum_i (x_{tr} - x_i, y_{tr} - y_i)$$

$$= \alpha[\sum_i (x_{tr} - x_i), \sum_i (y_{tr} - y_i)] = \alpha(nx_{tr} - \sum_i x_i, ny_{tr} - \sum_i y_i)$$

$$= \alpha n\left(x_{tr} - \sum_i \frac{x_i}{n}, y_{tr} - \sum_i \frac{y_i}{n}\right).$$

We may divide by αn, because that will not affect the direction of the vector. This yields the following:

$$\left(x_{tr} - \sum_i \frac{x_i}{n}, y_{tr} - \sum_i \frac{y_i}{n}\right).$$

The direction of this vector is the center-of-mass orientation, because the landmark's center of mass is at $(\Sigma_i x_i/n, \Sigma_i y_i/n)$.

Appendix B

Mean Acceptability Ratings for Each Placement of the Trajector for Each Spatial Relation for Wide and Tall Landmarks in Experiment 1

Tall					Wide				
				Above					
6.7	7.4	8.9	7.4	6.8	6.5	7.3	8.9	7.0	6.9
5.6	6.6	8.9	6.2	6.0	6.2	6.4	8.4	6.9	6.2
0.9	0.9	—	1.0	1.3	0.7	0.8	—	0.7	0.8
0.6	0.3	0.6	0.4	0.6	0.4	0.5	0.3	0.4	0.3
0.4	0.4	0.3	0.6	0.3	0.4	0.4	0.4	0.3	0.3
				Below					
0.4	0.5	0.3	0.3	0.3	0.5	0.6	0.2	0.4	0.9
0.3	0.3	0.4	0.4	0.3	0.5	0.8	0.6	0.3	0.4
0.8	1.0	—	0.9	0.9	0.8	0.8	—	0.7	0.9
5.9	6.2	8.6	6.4	5.4	5.9	6.8	8.8	6.4	6.1
6.4	7.2	8.6	6.9	6.6	6.8	7.2	8.8	7.5	6.5
				Left					
6.7	5.8	0.6	0.4	0.3	6.1	5.5	0.9	0.7	0.7
6.9	6.5	0.9	0.3	0.2	7.0	6.4	0.8	0.5	0.3
8.4	8.6	—	0.4	0.4	8.7	8.9	—	0.4	0.4
7.0	6.4	0.8	0.6	0.3	7.1	6.3	0.8	0.3	0.3
6.8	6.0	0.8	0.3	0.3	6.3	5.5	0.8	0.3	0.3
				Right					
0.6	0.3	0.8	5.9	6.9	0.4	0.5	0.9	6.0	6.7
0.5	0.7	0.8	6.9	6.9	0.4	0.5	0.9	6.3	6.6
0.4	0.3	—	8.7	8.6	0.4	0.4	—	8.6	8.8
0.3	0.3	1.2	6.8	7.2	0.5	0.3	1.0	6.3	6.9
0.4	0.2	0.7	6.0	6.7	0.3	0.6	1.1	5.4	6.2

Note. Dashes indicate the position of the reference object.

(*Appendixes continue*)

Appendix C

Mean Acceptability Ratings for Each Placement of the Trajector for Each Spatial Relation for Wide and Tall Landmarks in Experiment 2

Tall					Wide				
				Above					
6.6	7.3	8.7	7.7	6.9	6.7	7.0	9.0	7.4	7.1
6.3	6.7	8.6	7.0	6.3	5.9	6.8	8.9	6.7	6.4
1.2	1.1	—	1.5	1.2	1.1	1.2	—	1.2	1.6
0.3	0.4	0.5	0.4	0.4	0.6	0.6	0.4	0.7	0.7
0.5	0.4	0.3	0.3	0.5	0.6	0.5	0.9	0.9	0.9
				Below					
0.9	0.6	0.3	0.7	0.6	0.4	0.6	0.9	1.0	0.8
0.6	0.6	0.3	0.4	0.8	0.8	0.3	0.6	0.8	0.3
1.1	1.3	—	1.2	1.2	1.4	1.4	—	1.3	1.0
5.8	6.8	8.8	6.9	6.0	5.7	6.6	9.0	6.6	5.1
6.5	7.1	7.9	7.6	6.8	6.3	6.9	8.8	6.8	6.3
				Left					
6.3	5.8	1.1	0.8	0.6	6.4	6.1	1.4	0.6	0.6
6.8	6.2	1.3	0.6	0.8	7.2	6.6	1.4	0.7	0.5
9.0	8.7	—	1.2	0.4	8.7	8.8	—	0.4	0.4
6.6	6.1	1.0	0.6	0.6	7.1	6.7	1.0	0.5	0.7
6.5	5.8	0.9	0.3	0.7	6.5	6.0	1.2	0.7	0.3
				Right					
0.7	0.4	1.2	5.8	6.5	0.6	0.6	1.2	5.9	6.3
0.5	0.3	1.1	6.7	7.1	0.8	0.7	0.8	6.7	7.2
0.3	0.6	—	9.0	8.9	0.4	0.4	—	8.9	8.9
0.7	0.5	0.9	6.3	6.6	0.6	0.5	1.3	6.6	7.2
0.3	0.5	1.1	5.8	6.4	0.9	1.4	1.8	5.8	6.6

Note. Dashes indicate the position of the reference object.

Received October 5, 1998
Revision received March 28, 2000
Accepted May 15, 2000 ■