A probabilistic multidimensional model of location information

W. Todd Maddox1, William Prinzmetal2, Richard B. Ivry2, and F. Gregory Ashby3

1 Estes Lab. Department of Psychology, Harvard University, Cambridge, MA 02138, U.S.A.
2 Now at Department of Psychology, Arizona State University, Tempe, AZ 85287, U.S.A.
3 University of California at Santa Barbara, California, U.S.A.

Received January 6, 1993/ Accepted June 15, 1993

Summary. A probabilistic multidimensional model of location discrimination is developed and applied to data from an experiment in which subjects are required to determine whether a briefly presented horizontal and vertical bar are touching. The proposed gap-detection model assumes that errors in perception are due to variability in the perceived location and/or in the perceived length of the bars. A series of gap-detection models that allow variability only in perceived location were rejected on the basis of likelihood-ratio tests of overall goodness of fit. However, when the models were modified to account for: (a) a compression of the distance perceived between the bars (Wolford, 1975), or (b) the bisection illusion (Künnapas, 1955), excellent absolute fits to the data were obtained. A pair of models that suggests that the horizontal/vertical illusion or a response bias was operative failed. Applications of the model to more conventional object-perception experiments (e.g., the illusory-conjunction experiment) are discussed.

Introduction

It has been proposed that the mechanisms for determining the identity of visual features are different from the mechanisms for determining their location (e.g., Ungerleider & Mishkin, 1982). These two hypothetical systems are thought to be responsible for a functional difference between determining the what and where of a visual stimulus (e.g., Sagi & Julesz, 1985). While the independence of what and where information continues to be debated in psychophysical studies (see Johnston & Pashler, 1990), there are situations in which the correct classification of a visual stimulus is dependent on the accurate location of the stimulus features (e.g., Cohen & Ivry, 1989; Prinzmetal, 1981; Prinzmetal & Keysar, 1989; Treisman & Schmidt, 1982).

Consider the stimulus in Figure 1a and a task in which the subject is asked to determine whether the two bars are touching. Because of noise in the perceptual system, the location of the bars will not always be perceived correctly (e.g., Klein & Levi, 1987; Levi & Klein, 1989), and so on a certain proportion of trials the subject will make an incorrect response. In Figure 1b, for example, the two bars are physically separate (represented by the solid lines). However, because of perceptual noise, the bars are perceived as touching (represented by the broken lines). One way to model this phenomenon is to assume that each time the bars are presented, they are perceived in a unique location. Over trials, a distribution of perceived locations for each bar results.

The idea of a distributional perceptual representation is not new; in fact, it forms the basis of signal-detection theory (e.g., Green & Swets, 1967). Much of the early work in signal detection assumed a unidimensional probability distribution. This was reasonable, since most of the stimuli under investigation varied along only one physical dimension (e.g., intensity). The bars in Figure 1, however, are located physically in a two-dimensional space and so a unidimensional perceptual representation of perceived location is inadequate. There has been increasing interest and progress in modeling multidimensional perceptual representations (Ashby, 1992a; Ashby & Lee, 1991; Graham, Kramer, & Yager, 1987; Klein, 1985; Tanner, 1956; Wickens, 1992; Wickens & Olzak, 1989). This article extends this line of research.

In addition to inducing positional uncertainty, perceptual noise can cause a subject to misperceive other stimulus attributes. For example, in Figure 1c the subject correctly perceives the location of each bar, but perceives the length of the horizontal bar to be greater than its actual length, so an incorrect "touching" response is made. Figure 1d shows a case in which both the location of each bar and the length of the horizontal bar are misperceived.

Experiments have been conducted in which classification errors may result from the misperception of target-fea-
of object perception (e.g., attention manipulations, isoluminant stimuli).

The gap-detection task

The gap-detection task was devised to test the viability of a model that assumes that classification errors are caused by variability in perceived feature location. The basic stimulus is depicted in Figure 1a. On each trial, two bars were presented, one aligned horizontally and one aligned vertically. The bars were arranged so that together they roughly formed a percept of a T in one of four orientations: upright, upside down, or sideways facing left or right. However, the center-to-center distance of the two bars was varied so that on half of the trials the two bars were touching and on the other half of the trials there was a gap between the bars. Subjects were instructed to judge whether the bars were touching or not touching.

Several strategies were employed to increase the number of errors. These included: (1) the use of short-exposure durations; (2) subjects being required to perform a secondary task (i.e., to identify colored letters located in the center of the visual field); (3) the stimulus being masked; and (4) a large number of possible locations being used for the pair of bars. Other investigators (e.g., Klein & Levi, 1987; Levi & Klein, 1989), however, have found location errors with a brief exposure, but without any dual task, mask, or large number of possible stimulus locations.

The gap-detection model

The goal of this article is to test whether data collected in the gap-detection task can be accounted for by a quantitative model that assumes that errors are due to variability in the localization and identification of the stimulus features. The model specifies two types of parameters: (1) those corresponding to variability in perceived location; and (2) those corresponding to variability in perceived length of the bars.

Variability in perceived location

Consider the stimulus presented in Figure 2a. The two bars will be referred to as the cap and the stem, cap referring to the bar that is bisected by the stem. Denote the physical location of the center of the bar by \( \mu_c = (\mu_{bc}, \mu_{vc}) \) and \( \mu_s = (\mu_{bs}, \mu_{vs}) \), for the cap and stem of the stimulus, respectively. The subscript \( h \) denotes the position along the horizontal axis, and the subscript \( v \) denote the position along the vertical axis. Denote the perceived location of the center of the cap and stem by \( x_c = (x_{bc}, x_{vc}) \) and \( x_s = (x_{bs}, x_{vs}) \), respectively. In Figure 2a, the physical locations are denoted by the solid bars and the perceived locations by the dashed bars. We assume that perceptual noise induces trial-by-trial variability in the perceived locations. Specifically, we assume that for each stimulus the distributions of perceived locations along the horizontal and vertical axes are normal. Taken together, the perceived location of each bar has a bivariate normal distribution, which has a three-dimensional bell-like structure. The height of the bell at
variability may be decreased by selective attention to a dimension. Or stimulus-exposure duration may affect variability in perceived location, with short-exposure durations being associated with large variances. A perceptual dependence between the locations perceived along the horizontal and vertical axes would be accounted for by the covariance parameter (Ashby, 1988; Ashby & Townsend, 1986; Garner & Morton, 1969).

**Variability in perceived identity**

In the present context, variability in perceived identity causes the perceived height and width of the bars to vary from trial to trial. Inclusion of these sources of variability adds 10 parameters to the model (5 per bar). Identity parameters may be necessary to account for any illusory effects in which the perception of the bars individually or in combination is not veridical with the actual stimulus display. For example, the horizontal/vertical illusion and the bisection illusion may distort the mean (average) perceived heights and/or widths of bars arranged as in Figure 1 (see Results and theoretical analyses).

The most general version of the model is not parsimonious for several reasons—not the least of these being the fact that many of the parameters, although psychologically meaningful, are not identifiable. The approach taken here is to begin with a very simple one-parameter model that makes various assumptions about perceptual processing, and to relax these assumptions only when the fits of the model warrant it.

Often the process of adding constraints is guided more by mathematical convenience than by empirical validity. In this article, however, we shall test the simplifying assumptions empirically. If the data do not support a certain assumption, it can be replaced by a more general model with greater empirical validity. This approach will allow us to test a number of assumptions about perceptual processing. For example, parameters can be added to represent different perceptual illusions to evaluate the importance of these phenomena on the gap-detection task. We can also compare different dimensional spaces that could be used to describe the stimuli. For example, the dimensions may define either an object-centered or an environment-centered frame of reference. Within each of these models, it will be of interest to compare the relative variability along the two axes (e.g., Yap, Levi, & Klein, 1987).

An underlying assumption of the whole family of gap-detection models is that the distribution of perceived locations is normal. Although normality is usually assumed for mathematical convenience (see Ashby, 1992a, for several examples), in this article empirical support for this assumption was obtained in a pilot study conducted by one of us (Prinzmetal, unpublished data). In this experiment subjects indicated the location of a stimulus presented briefly. On each trial, subjects specified the location of a small gray circle by moving a mouse pointer to its perceived location. The stimulus could appear at any location along two imaginary circles, 2.29° or 4.58° from fixation. Three subjects participated in the experiment and completed approximately 3,600 experimental trials (1,800 at each eccentricity). The
distribution of perceived minus actual locations was computed for each subject. A bivariate normal distribution was then fit to this observed distribution for each subject by maximization of the correlation between the predicted and the observed distribution likelihoods. For all three subjects the correlations were highly significant (R > .95), suggesting that the normal distribution provides a good model of the subject’s perceived location judgments. It should be noted that other distributions (especially those that are unimodal and symmetric, such as the logistic), might also adequately describe the data.

Model computations

To model data from the gap-detection task, it is necessary to estimate the probability that the subject will respond “touching” on each trial. Consider the stimulus presented in Figure 2b. Note that the bars will touch physically if the distance between their centers along either the horizontal or the vertical axis is less than

\[ \frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} = \frac{l + w}{2} \]

where \( l \) is the length of each bar and \( w \) is the width. Now along the horizontal axis the distance between the centers is \( |\mu_{hc} - \mu_{hs}| \) and along the vertical axis the distance is \( |\mu_{vc} - \mu_{vs}| \). Therefore, the bars are physically touching if

\[ \max \left[ |\mu_{hc} - \mu_{hs}|, |\mu_{vc} - \mu_{vs}| \right] < \frac{(l + w)}{2} \]

(1)

Because of perceptual noise, the perceived location of each bar differs from trial to trial. Let the random vectors \( x_c = (x_{hc}, x_{vc}) \) and \( x_s = (x_{hs}, x_{vs}) \) represent the perceived center of each bar and suppose that \( x_c \) and \( x_s \) have bivariate normal (BVN) distributions with mean vectors \( \mu_c \) and \( \mu_s \) and covariance matrices \( \Sigma_c \) and \( \Sigma_s \), respectively. In order to decide whether the bars are touching, the subject is assumed to compute \( |x_{hc} - x_{hs}| \) and \( |x_{vc} - x_{vs}| \) and compare the larger of these two quantities with \( (l + w)/2 \). On any given trial, then, the bars are perceived to be touching when

\[ \max \left[ |x_{hc} - x_{hs}|, |x_{vc} - x_{vs}| \right] < \frac{(l + w)}{2} \]

(2)

Note that this decision rule is equivalent to the optimal classifier (i.e., the hypothetical device that maximizes accuracy: Ashby & Gott, 1988). The probability that the subject responds “touching” is equal to

\[ P(R_{Touching}) = P(\max \left[ |x_{hc} - x_{hs}|, |x_{vc} - x_{vs}| \right] < \frac{(l + w)}{2} ) \]

(3)

To compute this probability, let \( y \) be the vector \( y = (y_h, y_v) = (x_{hc} - x_{hs}, x_{vc} - x_{vs}) = (x_c - x_s) \)

1 In the most general version of the gap-detection model, \((l + w)/2\) would be a bivariate normally distributed random variable with a mean \(= \text{true width}/2 \). Excellent fits were obtained without any variability being assumed in the perceived length and width of the bars, and so these parameters were excluded from the model description. In some cases, however, the mean perceived length or width was included as a free parameter (see Results and theoretical analyses section entitled The perceptual-illusion hypothesis).

Note that \( y \) has a bivariate normal distribution with mean \( \mu_y = \mu_c - \mu_s \) and covariance matrix \( \Sigma + \Sigma \). Letting \( r = (l + w)/2 \), yields

\[ P(R_{Touching}) = P(\max \left[ |y_h|, |y_v| \right] < r) \]

where \( \text{bvn} \left( \mu_y, \Sigma \right) \) is the bivariate normal probability density function with mean \( \mu \) and covariance matrix \( \Sigma \). It is Equation 3 that we will need to evaluate. Appendix A provides the details of the model fitting procedure.

Method

On each trial subjects were briefly shown a display similar to that shown in Figure 3. The primary task was to indicate whether the two bars were touching. In order to make the gap-detection task more difficult, subjects also had to indicate the color of one of two letters located near the fixation point.

Subjects. Five students at the University of California, Santa Barbara, were tested. They were paid $6 per session.

Apparatus and stimuli. The stimuli were presented on a Zenith monitor (ZCM-1490) controlled by an AST Premium/286 computer. Each display contained four objects: two bars and two colored letters. The length and width of the bars spanned 32 and 8 pixels, respectively. At a viewing distance of 100 cm, these dimensions correspond to visual angles of approximately 47° and 12° of arc. One of the bars was oriented horizontally and the other was oriented vertically. Both bars were a moderately gray. One of the letters was an X and the other was an O. The letters spanned approximately 19° X 27° of arc. One letter was red and the other was green. In addition to the bars and colored letters, a fixation mark was used to begin each trial. This mark was a light-gray circle, approximately 23° in diameter. The four objects were presented on a black background.

The stimuli were arranged with the following constraints. The fixation point was located at the center of the monitor. The two bars were located on one of four sides of an invisible square surrounding the fixation point. The square is depicted by dotted lines in Figure 3. The length of each side of the square was 5.9'. Thus, the closest the center of a bar could be to the fixation point was just under 3'. If located in one of the corners, a bar was centered 4.2' from the fixation mark. The bars were placed on each of the four sides equally often. Once the side for a given trial was determined, the location on that side for one of the bars was chosen randomly. The location of the other bar was then determined by the spacing between the bars for that trial. The colored letters were presented at two of the possible locations near the fixation point. The center of these locations was displaced 27' up, down, left, or right from the fixation mark. The pairing of letters (X and O) to colors (green and red) was random.

Center-to-center distances of 16, 20, 24, and 28 pixels were used. The bars are touching for the two shortest distances, overlapping by four pixels for the shortest distance and just abutting for the other. There is a gap between the bars for the two farthest distances, spanning either 4 or 8 pixels. In visual angles, the gaps are either 5.3' or 11' or arc. The factorial combination of the four distances with four sides created 16 conditions for the gap-detection task: four sides on which the bars were located and four distances between the bars. Note that when the bars were to the left or to the right of the fixation mark, they formed either an upright or an upside-down T (although the cap and stem were not connected for the two farthest spacings). When the bars were above or below the fixation mark, they formed a sideways T, with the cap being either to the left or to the right of the stem.
Procedure. Subjects were seated approximately 100 cm in front of the computer monitor in a dimly lit room. The beginning of a trial was signaled by the appearance of the fixation mark. One second later, the fixation mark was replaced by the display of two colored letters and the pair of bars. After a variable exposure duration, the display was masked by an achromatic bright mask that extended 27° beyond the edges of the invisible square. The mask was visible for 1 s and then followed by a blank intertrial interval during which the subjects responded.

Responses were entered on the computer keyboard. The subject first made a judgment on the gap-detection task by typing 1 if the bars were touching or 0 if there was a gap between the bars. They then typed 1 if they thought the letter X was green or 0 if they thought the color of the X was red. No response was made concerning the color of the O. Error responses were indicated by auditory feedback. A 1,000-Hz tone indicated an error on the gap task. Two 500-Hz tones indicated an error on the colored-letter task. The subjects were instructed that, while both tasks were important, they should keep their eyes focused on the fixation mark to insure that they perceived the color of the X correctly.

The first day of testing was used to determine the appropriate exposure duration for the display. The exposure duration was initially set at 300 ms and subjects completed a block of 16 trials (4 sides × 4 spacings). Three more blocks of 16 trials were then run, with exposure durations of 214 ms, 171 ms, and 143 ms. At this point, the exposure durations were individually adjusted on the basis of the subject's performance. If the subject made zero or one error on the colored-letter task, the exposure duration was lowered by 1 raster cycle (70 Hz corresponding to 14.28 ms). If the subject made more than two errors on the colored-letter judgment, the exposure duration was increased by 1 raster cycle. No change was made when two errors were made on the colored-letter task. This procedure was repeated for four more blocks. After these seven short blocks, the subjects completed two blocks of 160 trials, each composed of 10 trials per condition. At the end of the first long block, the exposure duration was adjusted on the basis of the following criterion on the colored-letter task: lower by 1 raster cycle for less than five errors; hold constant for 5–10 errors; increase by 1 raster for more than 10 errors.

On subsequent days, subjects began testing with two practice blocks of 16 trials each, using the exposure duration from the last long block of the preceding session. They then completed three test blocks of 160 trials each. The exposure duration was adjusted after the first and second test blocks according to the criterion described above. The only exception was that the lowest exposure duration used was 29 ms.

Four of the subjects completed four test sessions or 12 blocks of 160 trials each and the fifth subject completed five sessions. The mean exposure duration on the first test block was 120 ms (sd = 30 ms; range = 100–171 ms). All of the subjects showed consistent improvement on the colored-letter task, so that by the last session the exposure duration was down to 29 ms for four of the subjects and 43 ms for the fifth subject (Subject 5).

Table 1. Frequency of “touching” and “not touching” responses for each subject by stimulus location (top/bottom, left/right) and center-to-center distance

<table>
<thead>
<tr>
<th>Dist</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rt</td>
<td>RNT</td>
<td>Rt</td>
<td>RNT</td>
<td>Rt</td>
</tr>
<tr>
<td>1</td>
<td>245</td>
<td>48</td>
<td>220</td>
<td>19</td>
<td>170</td>
</tr>
<tr>
<td>2</td>
<td>235</td>
<td>56</td>
<td>213</td>
<td>24</td>
<td>179</td>
</tr>
<tr>
<td>3</td>
<td>104</td>
<td>186</td>
<td>67</td>
<td>173</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>284</td>
<td>1</td>
<td>239</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>264</td>
<td>31</td>
<td>220</td>
<td>16</td>
<td>203</td>
</tr>
<tr>
<td>2</td>
<td>262</td>
<td>31</td>
<td>210</td>
<td>29</td>
<td>197</td>
</tr>
<tr>
<td>3</td>
<td>83</td>
<td>209</td>
<td>57</td>
<td>183</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>288</td>
<td>1</td>
<td>237</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: Rt = the observed frequency of “touching” responses; RNT = the observed frequency of “not touching” responses.

Results and theoretical analysis

Raw data

Subjects were correct on the colored-letter task on 98.4% of the trials. The data for the remaining 1.6% of the trials were not included in the subsequent analyses.

There were no consistent differences in performance on the gap-detection task between the pairs of conditions in which the bars appeared on opposing sides of the invisible square. The mean percentage correct on this task when the bars were on the left and right sides were 86.8% and 87.1%, respectively. The comparable percentages for the top and bottom were 81.9% and 84.2%. So the data were collapsed into eight conditions by combining across these pairs. There were over 1,200 observations per condition, or approximately 240 per subject.

Table 1 presents the raw frequencies of responding “touching” and “not touching” for the eight conditions by subject. A “touching” response is correct for the two smallest distances and incorrect for the two largest distances. A distance effect is apparent in the data. Subjects were much more likely to report that the bars were touching when the bars were separated by 5.5' (center-to-center distance 24 pixels, see Table 1) than when they were separated by 11' (center-to-center distance 28 pixels).

Subjects 1 and 4 made fewer errors when the bars were located on the left or right than in trials in which the bars were located in the top or bottom. For Subject 2 stimulus location had little effect. Subjects 3 and 5 were more accurate when the stimulus was on the left or right, but only when the bars were touching.

Parameter estimation and model testing

When the validity of a model is tested with respect to a particular data set, two problems must be considered. The first is to determine how unknown parameter will be estimated, the second is to determine how well the model describes (or fits) the data. The method of maximum likelihood is probably the most powerful method (see Ashby,
1992b; Wickens, 1982), and is especially useful for comparing the performance of nested models. In this article, nested as well as non-nested, models are compared. We therefore selected a goodness-of-fit statistic, Akaike’s (1974) an information criterion (AIC), which generalizes the method of maximum likelihood and allows for the comparison of models that are not nested. The AIC statistic is defined as

$$AIC(M_i) = -2\ln L_i + 2N_i$$  

(4)

where $N_i$ is the number of free parameters in Model $M_i$ and $\ln L_i$ is the log likelihood of the data as predicted by Model $i$ after its free parameter have been estimated via maximum likelihood. By the inclusion of a term that accounts for the number of parameters, a comparison can be made across models with different numbers of free parameters. When models are compared, the best is the one with the smallest AIC (see Sakamoto, Ishiguro, & Kitagawa, 1986, for a more thorough discussion of the minimum AIC procedure). All parameter estimates reported in this article were obtained with an iterative search procedure that minimized the AIC statistic.

**Modeling variability in perceived location**

The theoretical analysis begins with a set of models that allow variability only in the perceived location of the bars. The length and width of the bars ($l$ and $w$ from Equation 2, see also Figure 2a) and the mean center-to-center distances (i.e., $|\mu_e - \mu_s|$ from Equation 1, see also Figure 2a) are assumed to be perceived veridically. Because the bivariate normal distributions fit to the data of Prinzmetal (unpublished data) suggested little, if any, correlation between perceived horizontal and vertical locations, the covariance term in the gap-detection models was assumed to be zero (i.e., $\text{cov}_{xv} = \text{cov}_{wx} = 0$). The models differ only in the number of free parameters associated with variability in perceived location.

A total of six models were tested. The models are summarized in Figure 4. The $\sigma$ represent estimates of the variability in perceived horizontal and vertical location of the stem and cap of the T. Model 1 postulates no difference in the variability for perceived horizontal and vertical locations (one parameter; $\sigma_1$). Three two-parameter models were tested (Models 2–4). Like Model 1, Model 2 postulates no difference in the variability for perceived horizontal and vertical locations, but the model allows the amount of perceived variability to differ for stimuli presented on the top or bottom (i.e., the sideways T; $\sigma_3$) and left or right (i.e., the upright T; $\sigma_2$). This model should be able to account for the decrement in performance observed for stimuli presented on the top or bottom by yielding larger estimates of perceived location variability for these stimuli than for those presented of the left or right. Models 3 and 4 differ in the type of dimensional spaces that the parameters represent. Model 3 corresponds to environment-centered dimensions, namely one parameter represents variability in perceived horizontal location ($\sigma_1$), and another represents variability in perceived vertical location ($\sigma_2$). Model 4 defines the parameters with respect to object-centered coordinates. One parameter corresponds to perceived location variability parallel to the stem (i.e., along the horizontal axis for stimuli presented on the top or bottom and along the vertical axis for stimuli presented on the left or right; $\sigma_1$), and another parameter corresponds to perceived location variability parallel to the cap (i.e., along the vertical axis for stimuli presented on the top or bottom and along the horizontal axis for stimuli presented on the left or right; $\sigma_2$). Model 5 defines the parameters with respect to object-centered coordinates as well; however, perceived-location variability parallel to the stem for stimuli presented on the top or bottom ($\sigma_1$) is allowed to differ from that for stimuli presented on the left or right ($\sigma_3$). Finally, Model 6 allows separate estimates of variability in perceived horizontal location ($\sigma_1$) and vertical location ($\sigma_2$) for stimuli presented on the top or bottom and horizontal location ($\sigma_3$) and vertical location ($\sigma_4$) for stimuli presented on the left or right (see Appendix A for details of the model-based analyses).

![Fig. 4. Model parameters for six versions of the gap-detection model](image)

---

2 Note that Model A is said to be nested within Model B, when the setting of some of the parameters of Model B to constants makes the models equivalent.
Table 2. Goodness-of-fit values (AIC) for the gap-detection models that postulate variability only in perceived location

<table>
<thead>
<tr>
<th>Model</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of free parameters</td>
</tr>
<tr>
<td>1</td>
<td>2544.75</td>
</tr>
<tr>
<td>2</td>
<td>2542.84</td>
</tr>
<tr>
<td>3</td>
<td>2543.28</td>
</tr>
<tr>
<td>4</td>
<td>2541.43</td>
</tr>
<tr>
<td>5</td>
<td>2539.14</td>
</tr>
<tr>
<td>6</td>
<td>2541.14</td>
</tr>
</tbody>
</table>

Note: The best-fitting model for each subject is in bold type.

perceive the physical overlap between the bars (call this the overlap hypothesis). A second hypothesis is that subjects view this condition as one in which the bars are touching (call this the touching hypothesis), but where the stem is shorter than in the other three conditions (i.e., the three conditions in which the bars are physically abutting or are separated). There was no physical cue to indicate the extent of overlap; the brightness of this area was identical to nonoverlapping regions of the bars. In order to test rigorously between the overlap and touching hypotheses, two variants of all of the models were always compared—one assuming the overlap hypothesis and the other the touching hypothesis. Because the best-fitting model assumed the touching hypothesis, we shall discuss only models of this type.

The minimum AIC values for each of the six gap-detection models (see Figure 4) are presented in Table 2. The best fit for each subject is in bold type. The predictions from the best-fitting model for a representative subject (Subject 1) are presented in Figure 5a. An examination of the model predictions shows clearly that the models do not capture important aspects of the data. In particular, they seriously underpredict the large proportion of “touching” responses for the two smallest center-to-center distances. In addition, likelihood-ratio tests of overall goodness of fit (see Wickens, 1982, chapter 6) rejected these models for all subjects.

Fig. 5a—c. Predicted and observed proportion of “touching” responses for the four center-to-center distances for Subject 1. The squares and circles represent the data observed for stimuli presented on the top-bottom and left-right, respectively. The solid and broken curves represent the predictions from the best-fitting model (See Tables 2 and 3) for stimuli presented on the top-bottom and left-right, respectively. The predictions are for: a: the best-fitting location-only gap-detection model; b: the inter-object perturbation gap-detection model; and c: the bisection-illusion gap-detection model.

3 We thank Frank van der Velde (personal communication, April 1993) for suggesting this hypothesis.
4 In the overlap condition, the physical length of the stem is 12 pixels and the physical center-to-center distance is 16 pixels (see Methods section for details). Under the touching hypothesis, the physical length is assumed to be 28 pixels and the physical center-to-center distance is assumed to be 18 pixels.
Augmented models

There are several possible explanations for the poor performance of these models. Three of the most likely possibilities are tested next. One hypothesis, the inter-object perturbation hypothesis, is that the length and width of the bars (\(w\) and \(w\) from Equation 2, see also Figure 2a) are perceived veridically, but the mean center-to-center distances (i.e., \([\mu_c - \mu_s]\) from Equation 1, see also Figure 2a) are not. For example, there is a tendency for the perceptual location of stimuli to be closer to fixation than they actually are and this tendency is proportional to eccentricity (Chastain, 1986; Woldoff, 1975; Woldoff & Shum, 1980). In addition, Coren and Gigrus (1980) found that display items that group together are perceived as closer together than they actually are. In our displays, these phenomena would result in a reduction in the center-to-center distances of the two bars.

A second hypothesis, the perceptual-illusion hypothesis, assumes the opposite. Specifically, the mean center-to-center distances are perceived veridically, but the lengths and widths of the bars are not. Although there are several ways of allowing the mean perceived length of the base or of the stem to differ from the actual length, two perceptual illusions are particularly relevant to the gap-detection task, and they both predict that the perceived lengths of the two bars will be unequal. First, the horizontal/vertical illusion predicts that vertical lines are perceived as longer than horizontal lines. Second, the bisection illusion predicts that a bisector (i.e., the stem of the T) is perceived as longer than the bisection line (i.e., the cap of the T; Klüppelbusch, 1955). If the horizontal/vertical illusion is operating in the gap-detection task, then the perceived length of the vertical bar should be greater than the perceived length of the horizontal bar, regardless of whether the vertical bar is a stem or a cap. On the other hand, if the bisection illusion is operating, then the perceived length of the stem should be greater than the perceived length of the gap, regardless of whether the stem is in the horizontal or in the vertical alignment.

The third hypothesis, the response-bias hypothesis, is that mean center-to-center distances and length and width of the bars are perceived veridically, but that subjects have a bias to respond "touching."

To test these three hypotheses, each of the 6 gap-detection models (described in Figure 4) was augmented by including one additional free parameter. Specifically, for each of the 6 gap-detection models described in Figure 4, a single additional parameter was included that allowed either: a) the mean perceived center-to-center distance to differ from the actual center-to-center distance (inter-object perturbation hypothesis), b) the mean perceived length of the cap or stem to differ from the actual length (perceptual-illusion hypothesis), or c) a response bias to exist (response-bias hypothesis).

### Table 3. Goodness-of-fit values for the four augmented gap-detection models

<table>
<thead>
<tr>
<th>Model</th>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-Object Perturbation</td>
<td></td>
<td>1798.2</td>
<td>1169.7</td>
<td>1487.2</td>
<td>1353.0</td>
<td>1533.6</td>
</tr>
<tr>
<td>Bisection Illusion</td>
<td></td>
<td>1796.6</td>
<td>1163.1</td>
<td>1491.0</td>
<td>1354.0</td>
<td>1526.1</td>
</tr>
<tr>
<td>Horizontal/Vertical Illusion</td>
<td></td>
<td>2095.5</td>
<td>1650.9</td>
<td>1622.8</td>
<td>1618.8</td>
<td>1657.1</td>
</tr>
<tr>
<td>Response Bias</td>
<td></td>
<td>1833.5</td>
<td>1190.8</td>
<td>1518.6</td>
<td>1368.2</td>
<td>1622.8</td>
</tr>
</tbody>
</table>

*Note: The best fitting model for each subject is in bold type.*

to differ, by a constant, from the true center-to-center distances. This constant (call it \(c\)) was left as a free parameter in the model.

### The perceptual-illusion hypothesis

The horizontal/vertical-illusion hypothesis was implemented by allowing the mean perceived length of the vertical bar to be a free parameter. The bisection-illusion hypothesis was implemented by allowing the mean perceived length of the bisector to be a free parameter.

### The response-bias hypothesis

The response-bias hypothesis was implemented by taking the predicted probability of responding "touching" from the basic gap-detection model (i.e., from Equation 3) and multiplying it by a response-bias parameter, \(\beta\). Note that when \(\beta = 1\), there is no response bias. When \(\beta > 1\), subjects are biased toward the response "touching."

Table 3 presents the goodness-of-fit measure for the best fitting augmented versions of the gap-detection model for each subject. For Subjects 1, 3, and 5 the augmented versions of model 6 provided the best account of the data. For Subject 2, the augmented versions of model 4 provided the best fit, and for Subject 4, the augmented versions of model 5 were superior. The hypothesis (whether inter-object perturbation, horizontal/vertical illusion, bisection illusion, or response bias) that was supported by the data is in bold type. Predictions of the best fitting inter-object perturbation and bisection-illusion models are presented in Figures 5b and 5c, respectively, for Subject 1.

Several interesting results emerge. First, for all subjects, either the inter-object perturbation or the bisection-illusion model provided the best account of the data. For two subjects, the inter-object perturbation model provided the best fit, and for the remaining three subjects the bisection-illusion model provided the best fit. However, for all subjects, the goodness-of-fit values for these two models were almost identical. Unfortunately, on the basis of the fit values alone, it is impossible to choose one model over the other. Future research will be needed to compare these models. Second, both the inter-object perturbation and bisection-illusion models could not be rejected on the basis of likelihood-ratio goodness-of-fit tests for Subject 1 (\(p > .10\)), Subject 2 (\(p > .50\)), Subject 3 (\(p > .10\)), and Subject 4 (\(p > .50\)), suggesting excellent quantitative accounts of the...
Table 4. Parameter estimates of the best-fitting inter-object perturbation and bisection-illusion models for each subject

<table>
<thead>
<tr>
<th>Subject</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>$\sigma_3$</th>
<th>$\sigma_4$</th>
<th>$\sigma_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.15</td>
<td>5.50</td>
<td>1.64</td>
<td>1.05</td>
<td>-2.80</td>
</tr>
<tr>
<td></td>
<td>1.81</td>
<td>9.69</td>
<td>1.18</td>
<td>8.56</td>
<td>38.38</td>
</tr>
<tr>
<td>2</td>
<td>1.49</td>
<td>1.49</td>
<td>0.91</td>
<td>0.91</td>
<td>-2.70</td>
</tr>
<tr>
<td></td>
<td>1.23</td>
<td>1.23</td>
<td>7.79</td>
<td>7.79</td>
<td>38.15</td>
</tr>
<tr>
<td>3</td>
<td>1.96</td>
<td>8.77</td>
<td>1.70</td>
<td>1.64</td>
<td>-2.56</td>
</tr>
<tr>
<td></td>
<td>1.61</td>
<td>12.11</td>
<td>1.26</td>
<td>9.11</td>
<td>38.20</td>
</tr>
<tr>
<td>4</td>
<td>1.42</td>
<td>9.89</td>
<td>0.69</td>
<td>9.89</td>
<td>-3.09</td>
</tr>
<tr>
<td></td>
<td>1.45</td>
<td>10.19</td>
<td>0.69</td>
<td>10.19</td>
<td>38.18</td>
</tr>
<tr>
<td>5</td>
<td>1.29</td>
<td>12.58</td>
<td>1.84</td>
<td>0.90</td>
<td>-3.62</td>
</tr>
<tr>
<td></td>
<td>1.55</td>
<td>12.70</td>
<td>1.81</td>
<td>6.19</td>
<td>38.64</td>
</tr>
</tbody>
</table>

Note. $\sigma_1$ and $\sigma_2$ represent the estimated standard deviation for perceived location parallel to the stem and cap, respectively. $c = \text{the constant change in the mean perceived center-to-center distance}$; $l = \text{the length of the bisector (i.e., the stem)}$. All values are in pixels. The top value in each cell is for the inter-object perturbation model, and the bottom value is for the bisection-illusion model.

data. Only for Subject 5 was the best-fitting model rejected, and even in this case, both models accounted for over 97% of the variance in the data.

Third, note that the response-bias model is competitive with the inter-object perturbation and bisection models, although it never outperformed either of these models. The performance of the horizontal/vertical-illusion model, on the other hand, is clearly inferior. The fact that the bisection-illusion model proved superior to the horizontal/vertical-illusion model is in accord with previous research that shows that the bisection illusion is three times larger than the horizontal/vertical illusion when the two illusions are pitted against each other (Künnapas, 1955). Thus, the modeling results are in agreement with a prediction derived from perceptual research on these illusions.

The parameter values for the best-fitting inter-object perturbation and bisection-illusion models are presented in Table 4. First, note that the $c$ parameter from the inter-object perturbation model is negative for all subjects ($c$ ranges from $-3.62$ to $-2.56$). This result supports the hypothesis that the center-to-center distances were perceived as smaller than the true values. Second, in all cases the estimated length of the stem (i.e., the bisector) from the bisection-illusion model is larger (38.15–38.64 pixels) than the true length (32 pixels). Thus, the model fits confirm the basic effect of the bisection illusion in that the bisected segment is perceived as longer than the bisected segment. Indeed, the parameter estimates suggest that the magnitude of the illusion is about 20%, a value that is in agreement with those reported in the literature (Künnapas, 1955).

Other aspects of the parameter estimates should be noted. First, when different, the estimates of perceived-location variability parallel to the cap tend to be much larger than those for perceived-location variability parallel to the stem, especially for stimuli presented on the top or bottom. Perceived-location variability along the direction parallel to the caps will have little effect on performance unless the shift is large. For example, if the bars are abutting and the perceived location of each bar is vertical along the direction parallel to the stem, the perceived center-to-center distance along the axis parallel to the cap would have to increase by 21 pixels to produce a percept of two bars that are not touching. In contrast, if the bars are abutting, a single-pixel increase in the center-to-center distance along the direction parallel to the stem would yield an erroneous percept. The fact that classification errors are more likely with small shifts in perceived location in the direction parallel to the stem might influence subjects to attend more heavily to that direction. Greater attention to the direction parallel to the stem of the T should decrease variability in perceived location in this direction (Ashby & Perrin, 1988), thus keeping errors to a minimum.5

Second, note that the estimated-location standard deviations for all subjects (except Subject 5) are larger (or equal) for stimuli presented on the top or bottom than for stimuli presented on the left or right. In support of this result, we have found (unpublished data) that comparable performance is achieved only when exposure durations are longer for stimuli located on the top or bottom, than for stimuli presented on the left or right. Furthermore, a number of studies have found that for a given eccentricity, performance is worse for stimuli located above and below the fixation point than to the left or right of the fixation point (e.g., Rijsdijk, Kroon, & van der Wildt, 1980; Yund, Efron, & Nichols, 1990). This difference in performance may be due to the fact that the density of cones declines at a faster rate along the vertical than along the horizontal meridian (Cucio, Sloan, Packard, Hendrickson, & Kalina, 1987).

General discussion

The goal of this article was to develop and test a quantitative model of location discrimination. Our starting premise was that errors in object identification may be due to variability in the extraction of location information. Several versions of the model were applied to data collected in an experiment in which the stimulus was composed of two bars and the subject's task was to determine whether the bars were touching. Models that allow variability only in perceived location provided an unsatisfactory account of the data. Generalizing these models to account for the effects of (1) inter-object perturbations or (2) the bisection illusion provided excellent accounts of the data. Both models were superior to models that incorporated the horizon-

---

5 Hypothesizing that less perceived-location variability along the direction parallel to the stem is due to greater attention along this direction might appear at odds with the finding that the bisector was perceived as longer than its physical length. It seems reasonable to assume that greater attention in the direction parallel to the stem might in fact lead to more vertical perception of the bisected length. However, there is no reason to assume that attending to the location of a stimulus should necessarily lead to more vertical perception of the identity (i.e., the attributes, such as length and width) of the stimulus.
tal/vertical illusion or a response bias. Unfortunately, the current task does not allow us to distinguish between these two competing hypotheses.

We conclude with four examples of how the current modeling approach can be applied to data from more traditional object-perception experiments. First, an important method of examining the processes involved in object perception has been the illusory conjunction task introduced by Treisman and Schmidt (1982). We are presently applying a model that embodies the same basic assumptions as the gap-detection model to data collected in an illusory conjunction task using colored-letter stimuli (Ashby, Ivry, Maddox, & Prinzmetal, 1990). On each trial, two colored letters are presented and the subject’s task is to respond with the color of a target letter. The model assumes that on each trial, subjects independently perceive the locations of the target letter, the distractor letter, the target color, and the distractor color. Because of perceptual noise, the perceived locations vary from trial to trial and are assumed to be bivariate-normally distributed. On each trial, the reported color of the target letter is assumed to be the color that is perceived as closest to the perceived target location. Because of variability in perceived location, the distractor color will occasionally be perceived as closer to the target letter, leading to an illusory conjunction. The model also includes parameters associated with the probability of correctly identifying the target letter, target color, distractor letter, and distractor color.

The model is fit to more than just the trials on which an illusory conjunction occurred. Indeed, the model generates predictions for all of the possible responses a subject can make in this task. For example, it accounts for the trials on which the subject reports the incorrect target letter, but the correct color (i.e., letter-feature errors), and trials on which the subject reports the incorrect target letter and the incorrect target color (i.e., letter-feature illusory conjunctions). The model provided excellent fits of the data, with the best-fitting model requiring separate parameters for stimuli presented on the top or bottom and left or right, a result that converges with those of the present study.

This approach represents an improvement upon less formal studies of illusory conjunctions in which estimates of illusory conjunctions have proven difficult and generally involve rather strong assumptions (e.g., Cohen & Ivry, 1989; Cohen & Rafael, 1991; Prinzmetal, Presti, & Posner, 1986). The model-based approach is also promising because unique estimates of psychologically meaningful parameters (e.g., perceived location and identity) can be obtained.

A second example demonstrates how the modeling approach can be used to examine constraints shown to be operative during object perception. The formation of illusory conjunctions (and presumably of correct percepts) is affected by attentional manipulations (e.g., the spreading of attention). Prinzmetal and Keysar (1989) found that reading horizontally aligned digits increased illusory conjunctions between horizontally oriented items, and that reading vertically aligned digits increased the number of illusory conjunctions between vertically oriented items. Cohen and Ivry (1989) found that for a given inter-object distance the number of illusory conjunctions increased when the objects were between two digits that were to be reported. One explanation of these effects is that the spreading of attention affects the variability in perceived location of the features (Prinzmetal & Keysar, 1989). The gap-detection model and task provide an excellent mechanism for investigating the effects of attention on localization and identification processes. If, for example, the digit were placed in the periphery (e.g., horizontally or vertically), as opposed to in the center (see Figure 4), subjects would be required to spread their attention. Comparison of perceived-location-variability estimates, in this condition, with those obtained when the digits are presented in the center allows a direct test of the hypothesis that attention affects perceived-location variability.

The third example demonstrates how the model-based approach can be used to investigate another constraint identified in the study of illusory conjunctions; namely, that errors are more likely between objects that are similar in shape or color (Ivry & Prinzmetal, 1991). In the modeling of other perceptual and representational phenomena, it has become popular to attempt to predict performance in one experimental task from performance in another experimental task. For example, this approach has been used quite successfully to predict data collected in similarity and categorization tasks from performance in an identification task (e.g., Ashby & Lee, 1991; Nosofsky, 1986). On the assumption that the same perceptual processes are necessary in the traditional identification and illusory-conjunction experiments, it should be possible to predict performance in an illusory-conjunction task from performance in an identification task. The basic idea would be to estimate the identity parameters from identification data. These parameters would then be fixed in modeling the illusory-conjunction data, with only the location parameters being free to vary.

Finally, the fourth example involves a different phenomenon that has been shown to influence object perception. There has been a great deal of interest in recent years on perception with isoluminant stimuli (e.g., Livingstone & Hubel, 1987). For example, when luminance information is minimized, stereopsis is impaired (Lu & Fender, 1972; but see Jordan, Geisler, & Bovik, 1990) and structure from motion becomes difficult (Livingston & Hubel, 1987). Many of these phenomena may be predicted by means of a model similar to the gap-detection model. Location information may be worse with isoluminant stimuli than with stimuli that involve luminance differences (e.g., Gregory, 1977; Morgan & Aiba, 1985; Troscianko, 1987). Small errors in location could cause false matches in the correspondence between eyes (for stereopsis) or over time (for structure from motion). If this hypothesis is correct, it should be possible to model failures of stereopsis or structure from motion precisely, by considering the perceived location of image features.

Acknowledgements. This research was supported in part by Public Health Service Grant (NS30256) to RFI and by National Science Foundation Grant (BSN588-94063) to FGA. We would like to thank Asher Cohen, Anne Treisman, and Tom Wickens for their helpful comments. We also thank A. H. C. van der Heijden, Frank Van der Veld, and an anonymous reviewer for helpful comments on an earlier version of this article.
References


Appendix A

Model computations

In order to fit the gap-detection models to the data, Equation 3 (from the text) must be evaluated. Letting $f(y_h, y_v) = \text{bvn}(\mu, \Sigma)$, Equation 3 becomes

$$P(\text{RTouching}) = \int_{-r}^{r} \int_{-r}^{r} f(y_h, y_v) \, dy_h \, dy_v \tag{A-1}$$

Since the perceived horizontal and vertical locations are assumed to be independent (i.e., $\text{cov}_{hv} = 0$), Equation A-1 can be rewritten as

$$P(\text{RTouching}) = \int_{-r}^{r} f(y_h) \, dy_h \int_{-r}^{r} f(y_v) \, dy_v \tag{A-2}$$

where $f(y_h)$ and $f(y_v)$ are univariate normally distributed. The univariate integrals in Equation A-2 can be evaluated quite accurately (and quickly) by the use of an approximation to the cumulative normal distribution function (e.g., Ashby, 1992b).

Perceived location-only models

For the basic gap-detection models (see section entitled Modeling variability in perceived location), $r = (l + w)/2$ where $l$ is the length of the stem, and $w$ is the width of the cap. The densities, $f(y_h)$ and $f(y_v)$, are univariate normally distributed with means equal to the true horizontal and vertical center-to-center distances, and variances determined by the version of the gap-detection model under investigation (see Figure 4).

The inter-object perturbation model

This model is fit in the same fashion as the perceived location-only models, except that the mean horizontal center-to-center distance (for top or bottom stimuli) and the mean vertical center-to-center distance (for left or right stimuli) have a constant, $c$, added to each.

The perceptual-illusion models

Both the bisection- and horizontal/vertical-illusion models assume the densities, $f(y_h)$ and $f(y_v)$, are univariate normally distributed, with means equal to the true horizontal and vertical center-to-center distances, and variances determined by the version of the gap-detection model under investigation (see Figure 4). For the bisection illusion models, $r = (l + w)/2$ except when integrating $dy_h$ (for top or bottom stimuli) and $dy_v$ (for left or right stimuli), where $r$ was left as a free parameter. For the horizontal/vertical-illusion models, $r = (l + w)/2$ except when integrating $dy_v$ (for both top or bottom and left/right stimuli), where $r$ was left as a free parameter.

The response-bias model

In fitting the response-bias version of the gap-detection model, the probability obtained from Equation A-2 was multiplied by a constant, $\beta$, to determine the predicted probability of responding “touching.” Note that when $\beta = 1$, no response bias exists. When $\beta > 1$ there is a bias to respond “touching,” and when $\beta < 1$, there is a bias to respond “not touching.”

Computation of the AIC statistic

To determine the AIC statistic for Model k, the predicted probability of responding “touching” was incorporated into the following equation

$$\text{AIC}_k = -2 \sum_{i=1}^{2} \sum_{j=1}^{4} f_{ij} \ln (p_{ijk}) + \sum_{i=1}^{2} \sum_{j=1}^{4} (n_{ij} - f_{ij}) \ln (1 - p_{ijk}) + 2N_k$$

where $i$ = the stimulus locations (top or bottom or left or right), $j$ = the center-to-center distance, $n_{ij}$ = the total number of stimulus presentation in condition $ij$, $f_{ij}$ = the total number of observed “touching” responses in condition $ij$, $p_{ijk}$ is the predicted probability of responding “touching” for Model k in condition $ij$, and $N_k$ = the number of free parameters in Model k. An iterative parameter-search routine was used to determine the parameter values that minimized the AIC statistic.