Overestimation of base-rate differences in complex perceptual categories

W. TODD MADDOX and COREY J. BOHIL
University of Texas, Austin, Texas

The optimality of multidimensional perceptual categorization performance was examined for several base-rate ratios, for both integral and separable dimension stimuli, and for complex category structures. In all cases, the optimal decision bound was highly nonlinear. Observers completed several experimental sessions, and all analyses were performed at the single-observer level using a series of nested models derived from decision-bound theory (Maddox, 1995; Maddox & Ashby, 1993). In every condition, all observers were found to be sensitive to the base-rate manipulations, but the majority of observers appeared to overestimate the base-rate difference. These findings converge with those for cases in which the optimal decision bound was linear (Maddox, 1995) and suggest that base-rates are learned in a similar fashion regardless of the complexity of the optimal decision bound. Possible explanations for the consistent overestimate of the base-rate difference are discussed. Several continuous-valued analogues of Kruschke's (1996) theory of base-rate learning with discrete-valued stimuli were tested. These models found some support, but in all cases were outperformed by a version of decision-bound theory that assumed accurate knowledge of the category structure and an overestimate of the base-rate difference.

Categorization is fundamental to human survival. Every day we make hundreds of categorization judgments, and in many cases are very accurate. For example, our ability to categorize speech sounds and handwritten characters is unmatched by even the most sophisticated machines. Thus, it is reasonable to suppose that in certain domains our categorization performance is very nearly optimal (Ashby & Maddox, 1998). A major goal of the present study was to examine the optimality of human categorization performance when the observer is faced with a complex categorization problem similar in spirit to speech sound or handwritten character categorization.

To study optimality, one must rigorously define it. Although optimality can be defined in many ways, a reasonable definition, and the one we choose, is performance that maximizes long-run accuracy (or long-run reward; Ashby, 1992a; Green & Swets, 1966; Maddox & Ashby, 1993; Morrison, 1967). When the costs and benefits associated with each categorization response bias the observer toward one response or the other, performance that maximizes long-run accuracy may not maximize long-run reward. In the present study, cost and benefits were not manipulated, so the strategy that maximized long-run reward also maximized long-run accuracy. Maddox and Bohil (1998) exam-

ined the effects of cost–benefit manipulations on categorization performance and found that observers were sensitive to this type of information.

The optimal classifier is a hypothetical device that integrates information in such a way as to maximize long-run accuracy. For present purposes, two sources of information are relevant to the optimal classifier. One source is information about the distribution of category exemplars— that is, information about the central tendency and exemplar variability in each category. A second source is information about the category base-rates—that is, information about the likelihood of occurrence of each category. The optimal classifier uses these two sources of information to construct a decision rule that it uses to classify category exemplars. This article focuses on a comparison of human behavior with that of the optimal classifier when category base-rates were manipulated and the category structures were complex.

Although much anecdotal evidence suggests that humans are accurate categorizers, and thus should be sensitive to base-rate information, many studies suggest that people underestimate or ignore base-rate information, leading to the claim that often humans show *base-rate neglect* (see, e.g., Balla, 1982; Balla, Elstein, & Gates, 1983; Casscells, Schoenberg, & Graboys, 1978; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974; however, see Wallsten, 1981; Weber, Bockenholt, Hilton, & Wallace, 1993). These studies used story problems in which information about the category exemplars and base-rates was presented explicitly as percentages or probabilities. Gigerenzer, Hell, and Blank (1988; see also Gigerenzer, 1996; Gigerenzer & Hoffrage, 1995; Gigerenzer & Murray, 1987; however, see Kahneman & Tversky, 1996) provided evidence that the

prevalence of base-rate neglect is strongly affected by the medium used to present the base-rate information. Specifically, when base-rate information was presented in a probability format, base-rate neglect was often observed. However, when the base-rate information was presented in a frequency format, the appropriate use of base-rate information improved dramatically.

Estes, Campbell, Hatsopoulos, and Hurwitz (1989; see also Edgell et al., 1996; Edgell & Hennessey, 1980; Friedman, Massaro, Kitzis, & Cohen, 1995; Gluck & Bower, 1988; Holyoak & Spellman, 1993; Kruschke, 1996; Medin & Edelson, 1988; Nosofsky, Kruschke, & McKinley, 1992; Spellman, 1993) noted that much of our knowledge of base-rates is learned implicitly through experience with category exemplars. These studies used a hypothetical medical diagnosis task in which the category exemplars were composed of binary-valued symptom patterns. A wide range of results was found in these studies. Medin and Edelson found an *inverse base-rate effect*, in which the low base-rate category was most often selected when the observer was presented with conflicting symptom information. Gluck and Bower found the traditional base-rate neglect phenomenon. Nosofsky et al. (1992; see also Kruschke, 1996) found base-rate neglect for some symptom patterns, accurate base-rate estimates for other symptom patterns, and in some cases even an apparent overestimation of the base-rate information.

Kruschke (1996) proposed a theory of base-rate learning that can account for the wide range of base-rate sensitivities observed in the extant literature. The theory assumes that (1) base-rate information is learned and is applied consistently, (2) the high base-rate category is learned before the low base-rate category, and (3) the high base-rate category is learned by its typical features, while the low base-rate category is learned by its distinctive features. Because Kruschke's theory focuses on "features," it is restricted to the domain of discrete-valued stimuli, whereas many natural categories are composed of continuous-valued stimuli. The present study investigated several continuous-valued analogues of Kruschke's theory.

The aforementioned studies do not lend themselves to the study of optimality because in most cases, only a small number of binary-valued stimuli were used. Under these conditions, any one of a large (often infinite) number of decision rules maximizes accuracy. Of course, a test of optimality was not the goal of these studies; rather, the goal was to compare and contrast several models of category learning. For present purposes, however, a different approach was needed. In particular, we needed an approach in which only one decision rule was optimal.

Some of the early signal detection studies and several early categorization studies utilized such an approach (e.g., Green & Swets, 1966; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Ulehla, 1966). In the categorization studies, the two categories were usually univariate normally distributed, and category base-rates were manipulated. Because univariate normal distributions overlap and contain (essen-

tially) an infinite number of continuous-valued exemplars, only one decision rule maximizes categorization accuracy. First, the optimal classifier uses category distribution information to construct the *optimal decision function*. As in signal detection theory, the optimal decision function is determined by the likelihood ratio and is given by $l(x) = f(x|B)/f(x|A)$, where $f(x|i)$ denotes the probability density function for Category i and defines the likelihood of perceptual effect x given Category i. Second, the optimal classifier uses category base-rate information to construct the *optimal decision criterion*. The optimal decision criterion β_0 is given by $\beta_0 = P(A)/P(B)$ where $P(i)$ denotes the Category i base-rate. The optimal classifier then uses $l(x)$ and β_0 to construct the *optimal decision rule*:

If $l(x) > \beta_0$, then respond "B";
otherwise respond "A."   (1)

Notice that when $P(A) = P(B)$, the base-rates are equal, and the optimal decision criterion β_0 is placed where the likelihood ratio is equal to 1. When Category A has a higher base-rate, the optimal decision criterion is shifted in such a way that the observer might respond "A" even though the likelihood ratio information favors Category B. For example, if $P(A) = .75$, and $P(B) = .25$, the optimal decision criterion $\beta_0 = 3$. In other words, a Category B response would be given only when the likelihood that the stimulus comes from Category B is at least three times larger than the likelihood that the stimulus comes from Category A. The typical finding across all these studies is that observers use a decision criterion that is more conservative than the optimal decision criterion; in this example, a decision criterion less than 3. This result has been termed "conservative cutoff placement." Conservative cutoff placement can be contrasted with a finding of extreme cutoff placement. Extreme cutoff placement results when the observer uses a decision criterion that is more extreme than the optimal decision criterion—in this example, a decision criterion greater than 3. A reasonable psychological interpretation of conservative cutoff placement is that the observer underestimates the base-rate difference. A reasonable psychological interpretation of extreme cutoff placement is that the observer overestimates the base-rate difference.

Although these studies offer an important first step in the study of human performance when base-rates are manipulated, Maddox (1995) discussed several reasons why this work does not offer a definitive test of the optimality of human categorization performance. First, and most importantly, few rigorous models were developed to investigate suboptimalities in responding. In general, only the optimal classifier was investigated. Second, in many of the studies, observers were not highly experienced, and data were averaged across observers. Averaging of data from inexperienced observers often changes the qualitative structure of the data and might lead to incorrect inferences about human performance (Ashby, Maddox, & Lee, 1994; Estes, 1950; Maddox & Ashby, 1996). Finally, category exemplars varied along only a single dimension. Most real-world stimuli are multidimensional. When the category

exemplars vary along only one dimension, the optimal decision function is very simple. When category exemplars are multidimensional, the optimal decision function can range in complexity from a simple linear function to a highly nonlinear function.

In a series of studies, Ashby and colleagues (Ashby & Gott, 1988; Ashby & Maddox, 1990, 1992; Maddox & Ashby, 1993) examined the optimality of human categorization performance when category base-rates were equal but the optimal decision function varied in complexity. In many cases, the observer used the optimal decision function. When the optimal decision function was not applied, a function of the same form was utilized (e.g., a linear function was utilized when the optimal decision function was linear). Maddox (1995; see also Maddox & Bohil, 1998) examined the effects of unequal category base-rates in situations in which the optimal decision function was linear. Maddox found that a large number of observers used the optimal decision function and optimal decision criterion. However, in many cases observers showed "extreme cutoff placement"—that is, they used a decision criterion that was more extreme than that predicted by the optimal classifier. Maddox argued that these observers overestimated the base-rate difference. This finding was unexpected given the consistent finding in the previous literature that observers tend to underestimate the base-rate difference. (Consideration of these discrepant results and possible explanations will be reserved for the Discussion section.)

The present study extended Maddox's work (1995; see also Maddox & Bohil, 1998) to cases in which the categories were complex and the optimal decision function was highly nonlinear. To anticipate, the results of the present study converge nicely with those for cases in which the categories were simple and the optimal decision function was linear (Maddox, 1995; Maddox & Bohil, 1998). Specifically, observers were found to be sensitive to the category structures, but tended to overestimate the base-rate difference.

Before turning to the details of the experiment, we briefly introduce decision-bound theory (and several model-based instantiations of the theory), which motivated the present research. More thorough discussions of the theory are available elsewhere (see, e.g., Ashby, 1992a; Ashby & Lee, 1991, 1993; Ashby & Maddox, 1993, 1994; Ashby & Perrin, 1988; Ashby & Townsend, 1986; Maddox & Ashby, 1993; Thomas, 1995).

Decision-Bound Theory

Decision-bound theory (also called general recognition theory; Ashby, 1992a; Ashby & Lee, 1991, 1993; Ashby & Maddox, 1993; Ashby & Townsend, 1986; Maddox & Ashby, 1993) assumes that the experienced observer uses the same strategy as the optimal classifier, but with less success due to inherent limitations in human perceptual and cognitive processing. Consider an experiment involving two categories, A and B, whose exemplars vary along two perceptual dimensions. The optimal classifier perfectly records the perceptual representation for each stimulus i,

denoted by the vector $\mathbf{x}_i = [x_{1i}\, x_{2i}]'$. The optimal classifier then computes the optimal decision function $h_o(\mathbf{x})$ and the optimal decision criterion C_o. The optimal decision rule is to compare this optimal decision function and decision criterion, and to respond "A" or "B" according to the following rule:

If $h_o(\mathbf{x}) > C_o$, then respond "B";
otherwise, respond "A."   (2)

The Equation 2 decision rule effectively partitions the perceptual space into two response regions. Percepts falling into one region [where $h_o(\mathbf{x}) < C_o$] elicit Response A. Percepts falling into the other region [where $h_o(\mathbf{x}) > C_o$] elicit Response B. The partition between the regions [where $h_o(\mathbf{x}) = C_o$] is called the "optimal decision bound." The optimal decision function $h_o(\mathbf{x})$, which is the natural log of the Category B to Category A likelihood ratio, depends on the category means, variances, and covariance. In other words, it depends on the exemplar dimensional values. The optimal decision criterion C_o is determined from the base-rates using the following equation:
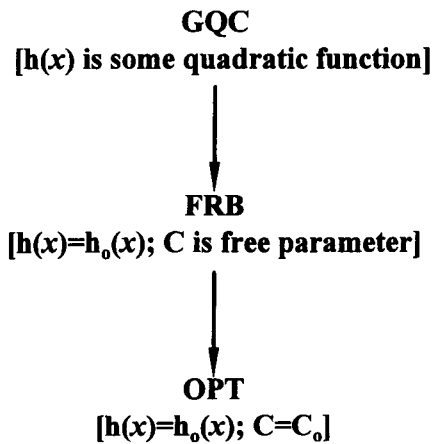
$$C_o = \ln [P(\mathrm{A})/P(\mathrm{B})].   (3)$$

Thus the optimal classifier must have knowledge of the category distributions and base-rate information in order to construct the optimal decision bound. Notice that the optimal decision rule in Equation 2 is equivalent to the decision rule in Equation 1 except that the natural log is applied to both sides of Equation 2. This one-to-one monotonic transformation does not affect the predictions of the optimal classifier. The natural log is applied strictly for practical purposes because it makes the effect of unequal base-rate ratios symmetric. For example, consider the following two unequal base-rate conditions: $P(\mathrm{A})$ to $P(\mathrm{B}) = .25$ to .75, and $P(\mathrm{A})$ to $P(\mathrm{B}) = .75$ to .25. In these cases, $\beta_o = \frac{1}{3}$ and 3, and $C_o = -1.099$ and 1.099, respectively. Taking advantage of this symmetry simplifies the explanation of the empirical results.

There are at least four limitations in human perceptual and cognitive processing that give rise to suboptimal responding. First, there is trial-by-trial variability in the perceptual information associated with each stimulus—that is, perceptual noise exists.[1] Let the vector $\mathbf{x}_i$ represent the observer's mean perceptual effect for stimulus i. Because of perceptual noise, the observer's percept of stimulus i, on any trial, is given by

$$\mathbf{x}_{pi} = \mathbf{x}_i + \mathbf{e}_p,   (4)$$

where $\mathbf{e}_p$ is a multivariate random variable (generally assumed to be normally distributed) that represents the effects of perceptual noise. Thus, each stimulus is represented perceptually by a multivariate normal distribution. Throughout this article, and in many applications of decision-bound theory, it is assumed that the perceptual covariance matrix $\Sigma_{pi} = \sigma_p^2 \mathbf{I}$. Under these assumptions, Σ_{pi} for each stimulus is identical, and is a scalar multiple of the identity matrix, $\mathbf{I}$. Because the perceptual covariance matrix is identical for each stimulus, this is called a stimulus-invariant

GQC
[h(x) is some quadratic function]

$\downarrow$

FRB
[h(x)=h_0(x); C is free parameter]

$\downarrow$

OPT
[h(x)=h_0(x); C=C_0]

Figure 1. Nested relationship among decision-bound models. The arrow points to a less general, "nested" model. GQC, general quadratic classifier; FRB, free base-rate optimal decision-bound model; OPT, optimal decision-bound model.

perceptual representation. Because the perceptual covariance matrix is a scalar multiple of the identity matrix, I, the perceptual variance along each dimension is identical, and perceptual independence holds. Clearly, in most cases these assumptions are incorrect (e.g., see Ashby & Lee, 1991; Maddox & Ashby, 1998). However, with high-contrast, response-terminated displays and fairly simple stimuli, as in the present study, this is often a reasonable assumption. With stimulus dimensions that yield complex perceptual representations, such as color patches that vary in hue, saturation, and brightness, or tones that vary in pitch, timbre, and loudness, these assumptions would be unsatisfactory.

Second, humans have imperfect knowledge of the category distribution parameters. Consequently, rather than using the optimal decision function $h_0(\mathbf{x}_{pi})$ to construct the optimal decision bound, observers use a suboptimal decision function $h_0(\mathbf{x}_{pi})$. However, decision-bound theory assumes that the observer uses a decision function of the same form as the optimal decision function (e.g., a linear function when the optimal function is linear). Third, observers often misestimate the base-rate information. Therefore, rather than using the optimal decision criterion C_0 to construct the optimal decision bound, observers use a suboptimal decision criterion, C, that is based on an estimate of the base-rates. Fourth, humans have imperfect trial-by-trial memory for the decision criterion (termed "criterial noise"). Criterial noise is represented by the random variable e_c, and is assumed to be univariate normally distributed with mean vector 0 and covariance matrix σ_c^2. With these limitations incorporated into Equation 2, decision-bound theory assumes that the experienced observer uses the following decision rule:

if $h(\mathbf{x}_{pi}) < C + e_c$, then respond "A";
otherwise, respond "B."   (5)

Thus, the probability of responding "A" on trials when stimulus i is presented is given by

$$P(A|\text{stimulus } i) = P\{[h(\mathbf{x}_{pi}) < C + e_c] |\text{stimulus } i\}. \quad (6)$$

Maddox (1995; see also Maddox & Ashby, 1993) developed several decision-bound models of categorization based on Equations 5 and 6 that differed in the form of the observer's decision function $h(\mathbf{x}_{pi})$ and the observer's decision criterion, C. These models included the general quadratic classifier, the optimal decision-bound model, and the free base-rate decision-bound model. These models are "nested" in the sense that some of the models can be derived from another model by setting some of the parameters of the more general model to constants. The nested structure of the models is depicted in Figure 1. Models at a higher level (toward the top of Figure 1) are more general than models at a lower level (toward the bottom of Figure 1).

The general quadratic classifier assumes that $h(\mathbf{x}_{pi})$ is some quadratic function. When the category covariance matrices are unequal, the optimal decision function is quadratic. The free base-rate decision-bound model assumes that the observer uses the optimal decision function [i.e., $h(\mathbf{x}_{pi}) = h_0(\mathbf{x}_{pi})$ from Equation 5], which in the present study was quadratic. This requires accurate knowledge of the category means, variances, and covariances, that determine $h_0(\mathbf{x}_{pi})$. However, the observer's knowledge of the base-rates is determined by estimating the value of C from their data. This model allowed us to determine whether the observer showed extreme or conservative cutoff placement. This model is nested under the general quadratic classifier because it assumes that the observer's decision function is the optimal (quadratic) decision function. Like the free base-rate decision-bound model, the optimal decision-bound model assumes that $h(\mathbf{x}_{pi}) = h_0(\mathbf{x}_{pi})$; however, instead of obtaining an estimate of the observer's base-rate knowledge, the optimal decision-bound model makes a specific assumption about base-rate knowledge, namely that the observer obtains an accurate estimate of the base-rates and assumes (correctly) $C = C_0$.

Each of the models depicted in Figure 1 was tested in the present study. One additional decision-bound model was developed to test the hypothesis that the observer shows base-rate neglect (i.e., that the observer assumes that the base-rates are equal). This model was overwhelmingly rejected by the data, suggesting that observers were indeed sensitive to base-rate information. In light of this fact, this model will not be considered further. Three additional models were proposed and each was tested in the present study. Each of these models represents a specific instantiation of Kruschke's (1996) theory. For ease of exposition these three models will be developed once the data are presented.

This article reports the results of a multidimensional perceptual categorization experiment in which the optimal decision bound was highly nonlinear. Five category A:B base-rate ratios (1:1, 1:2, 1:3, 2:1, and 3:1) were factorially combined with two sets of continuous-valued stimulus dimensions (one integral set and one separable set) for a total of 10 experimental conditions. The integral stimuli were rectangles that varied in height and width (Felfoldy,
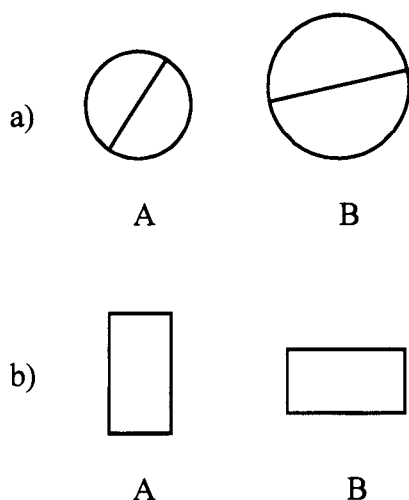
Figure 2. Representative stimuli from (a) Condition C (circular stimuli) and (b) Condition R (rectangular stimuli).

1974). The separable stimuli were circles with a radial line that varied in circle size and line orientation (Felfoldy & Garner, 1971; Garner & Felfoldy, 1970). The optimal decision bound in each of the five base-rate conditions was nonlinear, but was unique to each condition (see Figure 3). In fact, as Figure 3 (to be discussed shortly) suggests, the specific base-rate ratio had a large effect on the shape of the optimal decision bound (cf. Figures 3b and 3c). Thus, this experiment examines base-rate knowledge across a wide range of nonlinear decision bounds for two sets of qualitatively different stimulus dimensions. In addition, the categories were chosen in such a way as to facilitate the examination of several continuous-valued analogues of Kruschke's (1996) theory of base-rate learning. In all experimental conditions, each observer participated in several experimental sessions to ensure that performance had reached (or nearly reached) asymptote. In addition, all model-based analyses were conducted at the individual observer level.

METHOD

Subjects

All observers were volunteers from the university community. All observers claimed to have 20/20 vision or vision corrected to 20/20. There were 6 observers in each of the four unequal base-rate conditions (3 for each stimulus type) and 8 observers in the equal base-rate condition (4 for each stimulus type) for a total of 32 observers. All observers completed four experimental sessions except for Observer 3 from the 2:1 base-rate, rectangle condition, and Observer 1 from the 3:1 base-rate, rectangle condition, who were unable to complete the final experimental session. Observers were paid a base salary of $5 per session (each lasting approximately 1 h) and were eligible for a daily bonus. Most observers received bonuses in the range of $1 to $3.

Stimuli and Stimulus Generation

Examples of the circles (C) and rectangles (R) are depicted in Figure 2. The stimuli were computer generated and displayed on an SVGA monitor in a dimly lit room. The stimuli were white on a black background and subtended approximately 1° of visual angle.

There were two categories, A and B, each defined by a specific bivariate normal distribution. The parameters for each category are displayed in Table 1. Prior to the experiment, a set of 300 stimuli were sampled randomly from the category distributions, with the constraint that the number drawn from each distribution reflected the population category base-rates for each of the five conditions (e.g., 150 from each distribution in the 1:1 base-rate condition; 200 from Category A and 100 from Category B in the 2:1 base-rate condition; 100 from Category A and 200 from Category B in the 1:2 base-rate condition; 225 from Category A and 75 from Category B in the 3:1 base-rate condition; 75 from Category A and 225 from Category B in the 1:3 base-rate condition). Each sample (defined by the ordered pair x_1, x_2) was used to construct a stimulus. A stimulus was constructed with either diameter (for circles) or width (for rectangles) equal to the x_1 coordinate and orientation (for circles) or height (for rectangles) equal to the x_2 coordinate. Because the dimensions of the circle stimuli are measured in different units, it was necessary to equate diameter and orientation units to determine the optimal bound. In an attempt to equate diameter and orientation discriminability, 1 unit of diameter was made equal to $\pi/1,000$ radians. For each condition, four random orderings of these stimuli were then constructed. Each random ordering was assigned to a particular session so that each observer in the same condition was presented with the same stimulus sequence during each session. This sampling and presentation procedure reduces error variance by ensuring that (1) the sample statistics were identical across sessions and (2) the presentation sequence was fixed across observers.

On the basis of the bivariate normally distributed category populations, the optimal classifier predicts 84%, 85%, 87%, 87%, and 90% accuracy for the 1:1, 2:1, 1:2, 3:1, and 1:3 base-rate conditions, respectively. A bivariate normal distribution looks like a three-dimensional bell-shaped structure in which the height of the bell at any point represents the likelihood that a random sample from the population will have the x_1 and x_2 values associated with that point. A convenient method of presenting the same information is via the contours of equal likelihood, each of which is created by taking a slice parallel to the (x_1, x_2) plane and looking down on the result from above. This is called a contour of equal likelihood because all points on the same contour are associated with the same likelihood (i.e., probability density). For bivariate normal distributions, the contours of equal likelihood are always elliptical. The height at which the slice is taken is arbitrary, and so a single contour sufficiently describes all the relevant information about the category.

Figure 3a depicts the elliptical contour of equal likelihood for each category used in the experiment. These represent the stimulus regions and are denoted by the labels A and B. Notice that the Category A contour is tilted clockwise from vertical, whereas the Category B contour is tilted counterclockwise from vertical. The contours are tilted because each category distribution has a nonzero covariance term. In Category A, the covariance term is positive, and in Category B, the covariance term is negative. Therefore, in Category A exemplars with large x_1 values tend to have large x_2 values (i.e., the x_1 and x_2 values covary, or are correlated). In Category B, on the other hand, large values of x_1 tend to be associated with small values of x_2. The magnitude of the correlation is .80 in both categories, the direction of the correlation is positive in Category A and is negative in Category B. (The correlation is computed by dividing the covariance by the product of the standard deviations.) Interestingly, though, the magnitude of the tilt appears larger in Category B than in Category A (see Figure 3a). In short, the fact that the x_2 variance is larger than the x_1 variance in Category A, whereas the x_1 and x_2 variances are equal in Category B, leads to a less extreme tilt in Category A.

Figure 3b depicts the optimal decision bound and response regions, "A" and "B" associated with the 1:1 base-rate condition. Figure 3c depicts the optimal decision bounds and response regions as-

Table 1
Parameter Values of the Category Distributions (in Pixels)

Parameter	Category	
	A	B
μ_x	270	300
μ_y	400	400
σ_x	52	65
σ_y	132	65
cov_{xy}	5,491	−3,401

Note—x, width (or diameter); y, height (or orientation); cov_{xy}, covariance between width and height (or diameter and orientation). When the y dimension represented orientation, the pixel units were converted to radians by multiplying the values in the table by the scaling factor $\pi/1,000$.

sociated with the 2:1 (solid hyperbolae) and 3:1 base-rate (broken hyperbolae) conditions. Figure 3d depicts the optimal decision bounds and response regions associated with the 1:2 (solid hyperbolae) and 1:3 base-rate (broken hyperbolae) conditions. The optimal decision function $h_o(x)$ is determined from the Category A and B means, variances, and covariance terms (see Equation 2). These parameters are provided in Table 1, and the resulting contours of equal likelihood are depicted in Figure 3a. The optimal decision function $h_o(x)$ satisfies $.000177395x_1^2 - .00025648x_2^2 - .000865x_1x_2 + .27035x_1 + .45495x_2 - 133.26375$. The optimal decision criterion C_o satisfies 0, .693, −.693, 1.099, and −1.099 in the 1:1, 2:1, 1:2, 3:1, and 1:3 conditions, respectively (see Equation 3). The optimal decision bounds depicted in Figures 3b–3d were constructed by subtracting the optimal decision criterion, C_o, from the optimal decision function, $h_o(x)$, and then setting the resulting equation equal to zero. Figure 4 displays the stimuli used in the 2:1 (panel a), 1:2 (panel b), 3:1 (panel c), and 1:3 (panel d) base-rate conditions. The plus signs denote stimuli from Category A, and the periods denote stimuli from Category B. The optimal decision bound is also included in each panel.

The optimal decision bound is defined in the physical space, where x_1 is the size of the circle (or width of the rectangle), and x_2 is the orientation of the radial line (or the height of the rectangle). An obvious issue is to determine how the optimal decision bound is defined in the perceptual space. If the purpose is to identify optimal responding, then a close correspondence between the physical space and the perceptual space is not required. The only requirement is that the transformation from the physical to the perceptual space is one-to-one (see Ashby & Maddox, 1992, note 3 for a proof). Under these conditions, optimal bounds in the physical space will manifest themselves as optimal bounds in the perceptual space. This property is likely satisfied with the circle and rectangle stimuli used in the present study. For example, the Stevens (1961) exponent for length is close to 1.0, so perceived and physical size should be nearly identical. In addition, psychological scaling solutions have been derived for the circular stimuli (Nosofsky, 1985, 1986, 1987; Nosofsky, Clark, & Shin, 1989; Shepard, 1964) and for the rectangular stimuli (Schonemann, 1977). In both cases, the resulting psychological space was very similar to the physical space.

Notice that the optimal bound and response regions in the 1:2 and 1:3 base-rate conditions (Figure 3d) look very similar to the optimal bound and response regions in the 1:1 base-rate condition (Figure 3b), whereas the optimal bound in the 2:1 and 3:1 base-rate conditions (Figure 3c) look different from the optimal bound in the 1:1 base-rate condition. All five bounds are related via a simple intercept shift, but the increased Category A base-rate in the 2:1 and 3:1 conditions leads to a change in the category label assigned to stimuli that fall in the middle region of the stimulus space, and thus to the shape of the optimal bound. (It is well established in analytic geometry that the shape of a hyperbola can vary depending on the value of the constant term.) When the base-rates are equal, Category B tends to dominate Category A because of the strong negative correlation (and thus counterclockwise tilt). (This large negative

correlation leads to a large eigenvalue along the major axis of the ellipse.) This domination increases as the base-rate for Category B is increased, thus yielding a decision bound and response regions of a similar shape. However, as the base-rate for Category A is increased a reversal occurs, and the shape of the optimal decision bound and the response regions change. Although the exact base-rate ratio has a large effect on the orientation of the optimal decision bound and could lead to qualitatively different levels of performance in the 2:1 and 3:1 base-rate conditions as compared with the 1:2 and 1:3 base-rate conditions, as we will see shortly, observers in all conditions behaved in a similar manner.

Procedure

The experiment was conducted in a dimly lit soundproof chamber. The observer was seated approximately 30 in. from the computer monitor with no head or chin restraints. Observers were told that perfect performance was impossible. However, a nearly optimal level of performance was specified as a goal (in the form of desired accuracy). Observers were instructed to maximize accuracy and not worry about speed of responding. A typical trial proceeded as follows: A stimulus was presented on the screen and remained until a response was made. The observer's task was to classify the presented stimulus as a member of Category A or B by pressing the appropriate button. Following a response, the screen went blank for a period of 500 msec. Next, corrective feedback was presented in the center of the screen for 500 msec. Finally, the screen went blank again for an intertrial interval of 500 msec, followed immediately by the next stimulus presentation. Each experimental session consisted of a 100-trial practice session, followed by a 300-trial experimental session. Observers were given periodic breaks.

RESULTS AND THEORETICAL ANALYSIS

Final Session Analyses

Accuracy rates. Each observer's final session accuracy is displayed in Table 2. The results can be summarized as follows. First, accuracy rates were quite high, although in every case they were suboptimal. Second, accuracy rates did not differ significantly across stimulus types [$t(30) = 0.95, p > .05$], 1:2 and 2:1 base-rate conditions [$t(10) = 1.04, p > .05$], or across 1:3 and 3:1 base-rate conditions [$t(10) = 0.91, p > .05$]. One additional analysis was conducted to determine whether performance in the 1:2 and 1:3 conditions differed from performance in the 2:1 and 3:1 conditions. Recall from Figure 3 that the stimuli in the middle region should be optimally assigned to Category B in the 1:2 and 1:3 base-rate conditions and to Category A in the 2:1 and 3:1 conditions. In order to determine whether this asymmetry affected performance, we computed the proportion of "high base-rate" responses for each observer in each condition. If the difference between the 1:2 and 2:1, and 1:3 and 3:1 optimal bounds was not adversely affecting performance, we would expect the proportion of high base-rate responses to be nearly identical across conditions. The proportions of high base-rate responses in the middle region were .769, .852, .896, and .880 in the 1:2, 2:1, 1:3, and 3:1 base-rate conditions, respectively; t tests between the 1:2 and 2:1 high base-rate proportions [$t(10) = 1.41, p > .05$] and between the 1:3 and 3:1 high base-rate proportions [$t(10) = 0.41, p > .05$] were nonsignificant. In light of these results, nearly all subsequent analyses are based on data collapsed across stimulus type and reciprocal base-rate conditions. In addition, the collapsed 1:2 and
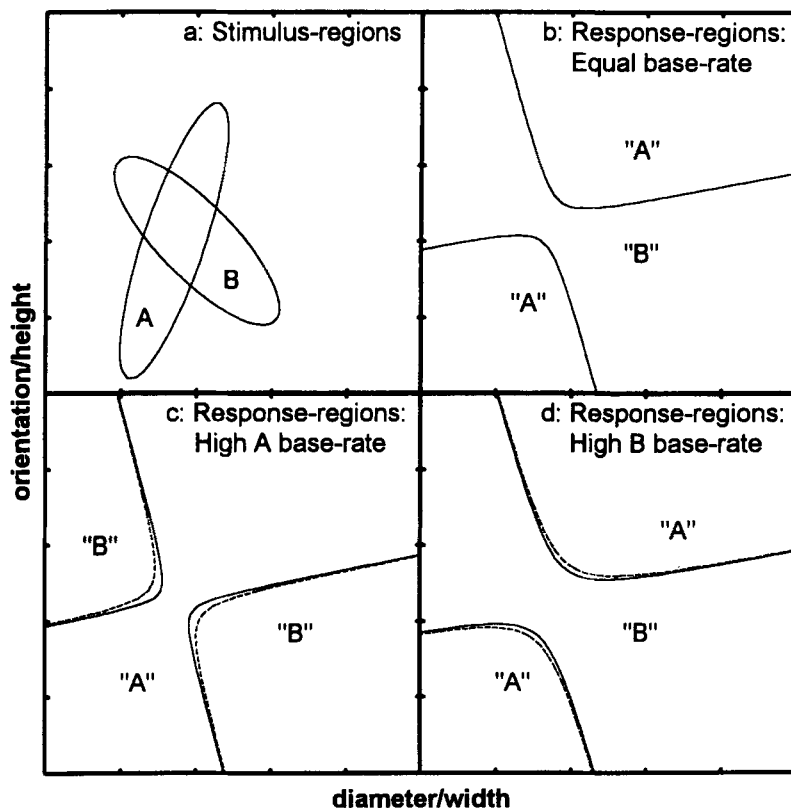
Figure 3. (a) Stimulus regions and elliptical contours of equal likelihood for Categories A and B, (b) response regions and optimal decision bound for the 1:1 base-rate condition, (c) response regions and optimal decision bounds for the 2:1 and 3:1 base-rate conditions, and (d) response regions and optimal decision bounds for the 1:2 and 1:3 base-rate conditions. In panels c and d, the optimal decision bound in the 1:2 and 2:1 conditions is represented by the solid curve, and the optimal decision bound in the 1:3 and 3:1 conditions is represented by the broken curve. The contours of equal likelihood in panel a are derived by setting the category likelihood to .000025.

2:1 base-rate condition data are referred to as the *moderate* base-rate condition, and the collapsed 1:3 and 3:1 base-rate condition data are referred to as the *large* base-rate condition.

Model-based analyses. The three decision-bound models depicted in Figure 1 were applied to each observer's final session data using a maximum likelihood parameter estimation procedure (for details see Ashby, 1992b; Maddox, 1995; Wickens, 1982). Consider an experiment with Categories A and B and a set of n stimuli, $S_1, S_2, \ldots, S_n$. For each stimulus, a particular model predicts the probabilities that the observer will respond "A" and "B," which we denote by $P(A|S_i)$ and $P(B|S_i)$, respectively. The results of an experimental session are a set of n responses, $r_1, r_2, \ldots, r_n$, where we arbitrarily set $r_i = 1$ if Response A was made to stimulus i and $r_i = 0$ if Response B was made. According to the model, and assuming the responses are independent, the likelihood of observing this set of n responses is

$$L(r_1, r_2, \ldots, r_n) = \prod_{i=1}^{n} P(A \mid S_i)^{r_i} P(B \mid S_i)^{1-r_i}.$$

The maximum likelihood estimators are those values of the unknown parameters that maximize $L(r_1, r_2, \ldots, r_n)$. For practical reasons, it is common to take the natural log of the likelihood function, multiply through by -1, and then minimize the $-\ln(L)$. It is this approach that was taken in this study.

The general quadratic classifier has seven free parameters: five decision-bound parameters, a perceptual noise parameter, and a criterial noise parameter. The free base-rate decision-bound model assumes the optimal decision function [i.e., $h(x) = h_0(x)$], but estimates the observer's base-rate knowledge from the data by leaving the decision criterion C (hereafter referred to as C_f) from Equation 5 as a free parameter. In addition, σ_p and σ_c are left as free parameters, for a total of three parameters. The optimal decision-bound model assumes the optimal decision function [i.e., $h(x) = h_0(x)$] and the optimal decision criterion (i.e., C = C_o), but includes two free parameters, σ_p and σ_c. In other words, this model postulates that the observer has accurate knowledge of the base-rate difference.

These models are "nested" in the sense that a more restricted model can be derived from a more general model
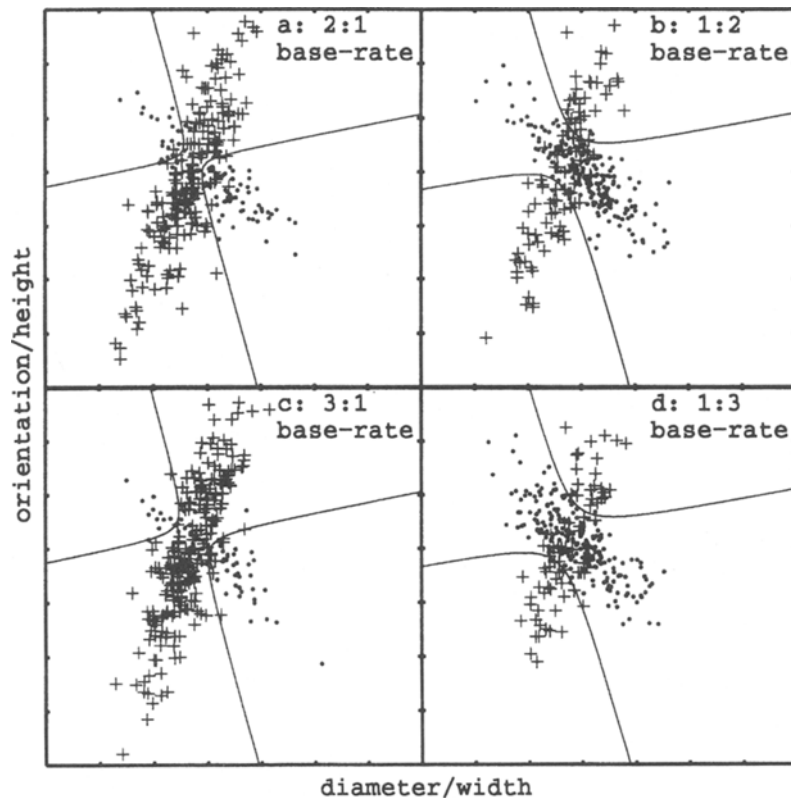
Figure 4. Stimuli used in the (a) 2:1 base-rate condition, (b) 1:2 base-rate condition, (c) 3:1 base-rate condition, and (d) 1:3 base-rate condition. In each plot, the plus signs denote Category A stimuli, and the points denote Category B stimuli. In addition, the solid curve represents the optimal decision bound.

by setting some of the parameters of the more general model to constants. For example, the free base-rate model is nested within the general quadratic classifier because the free base-rate model can be derived from the general quadratic classifier by setting the coefficients of the quadratic function to the coefficients of the optimal decision function. Similarly, the optimal decision-bound model can be derived from the free base-rate model by setting the decision criterion to the optimal value, C_o. Because these models are nested, it is always the case that the general quadratic classifier will provide the best account of the data, and the optimal decision-bound model will provide the worst account of the data. Even so, because these models are nested, a series of G^2 tests can be performed to determine the most parsimonious model—that is, the model with the fewest number of free parameters that was not "significantly" improved upon by a more general model (Ashby, 1992b; Wickens, 1982). A series of G^2 tests was performed in this study. In order to perform these G^2 tests, we had to choose a significance level, α. We performed our analyses twice, once with $\alpha = .05$ and once with $\alpha = .01$. For 24 of the 32 observers, the "most parsimonious model" was not affected by the value of α. For 8 of the 32 observers, the value of α did affect the identity of the most parsimonious model. For these 8 observers, it was always

the case that $\alpha = .05$ led to the selection of a more general model as the most parsimonious. This must be the case because the value of α determines the amount of evidence against the null hypothesis that is necessary to reject the null hypothesis. The smaller the value of α, the more evidence is needed against the null hypothesis in order to reject it. Because the null hypothesis assumes that the general model does not provide a significant improvement over the less general model, it follows that the smaller value of α will lead to more cases in which the less general model is selected.

Within the framework of decision-bound theory, it is assumed that the experienced observer attempts to respond optimally. Thus, a decision-bound theorist expects that the less general model (e.g., the optimal decision-bound model) will provide the most parsimonious account of the data. As α decreases, more evidence is needed to reject the less general model, so the smaller the value of α, the more evidence will be garnered in support of decision-bound theory. In light of this fact, we report the analyses based on $\alpha = .05$. Even so, none of our general conclusions were affected by the choice of α.

Base-rate neglect. As stated earlier, the hypothesis that observers are insensitive to base-rate differences was overwhelmingly rejected by the data.

Table 2
Accuracy Rates and Goodness-of-Fit ($-\ln L$) Values for Final Session of Each Observer

O	% Correct	OPT (2 Parameters)	FRB (3 Parameters)	GQC (7 Parameters)	% of Responses Accounted for by the Most Parsimonious Model
			1:1		
C/1	82	110.394	109.919	**103.227**	86
C/2	79	**119.571**	119.083	115.939	82
C/3	84	62.335	56.096	**49.557**	91
C/4	80	112.548	**108.207**	108.193	84
Mean	81	101.212	98.326	94.229	86
R/1	72	139.608	139.522	**117.163**	81
R/2	79	94.119	79.631	**71.789**	89
R/3	80	**99.928**	99.056	98.856	82
R/4	83	**80.469**	80.275	80.117	88
Mean	79	103.531	99.621	91.981	85
Overall 1:1 mean	80	102.372	98.974	93.105	85
			1:2		
C/1	83	77.561	**75.177**	71.996	88
C/2	82	83.451	**63.706**	59.308	93
C/3	77	158.592	151.164	**142.217**	79
Mean	81	106.535	96.682	91.174	87
R/1	85	**78.301**	77.252	76.084	90
R/2	85	**69.786**	69.657	68.027	89
R/3	83	**92.669**	92.414	90.652	86
Mean	84	80.252	79.774	78.254	88
Overall 1:2 mean	82	93.393	88.228	84.714	88
			2:1		
C/1	79	110.512	**97.415**	94.354	87
C/2	73	134.190	95.700	**80.127**	89
C/3	84	83.168	**80.282**	80.237	87
Mean	79	109.290	91.132	84.906	88
R/1	85	**74.201**	73.585	73.268	92
R/2	76	126.892	96.213	**85.202**	91
R/3	82	115.489	107.016	**89.511**	85
Mean	81	105.527	92.271	82.660	89
Overall 2:1 mean	80	107.409	91.702	83.783	89
Moderate mean	81	100.401	89.965	84.249	88
			1:3		
C/1	80	129.995	68.400	**62.511**	91
C/2	86	66.927	66.927	**53.484**	93
C/3	76	151.112	**90.425**	87.262	89
Mean	81	116.011	75.250	67.752	91
R/1	81	116.090	**103.895**	103.119	88
R/2	86	67.444	67.080	**60.781**	92
R/3	89	43.043	37.328	**30.238**	96
Mean	85	75.526	69.434	64.713	92
Overall 1:3 mean	83	95.768	72.342	66.232	92
			3:1		
C/1	86	74.948	67.550	**60.097**	92
C/2	83	85.116	82.588	**65.483**	91
C/3	85	40.002	36.891	**30.271**	96
Mean	85	66.689	62.343	51.950	93
R/1	85	64.219	**55.756**	55.560	93
R/2	84	56.274	**51.558**	46.887	94
R/3	86	**59.623**	59.359	57.845	93
Mean	85	60.039	55.558	53.431	93
Overall 3:1 mean	85	63.364	58.950	52.691	93
Large mean	84	79.566	65.646	59.462	92

Note—The "most parsimonious" model for each observer is in bold type. O, observer; OPT, optimal decision-bound model; FRB, free base-rate optimal decision-bound model; GQC, general quadratic classifier; C, circle stimuli condition; R, rectangle stimuli condition. Moderate mean refers to the moderate base-rate ratios (1:2 and 2:1). Large mean refers to the large base-rate ratios (1:3 and 3:1). The final session was the third session for 2 observers and the fourth session for the other 30 observers.

Optimal base-rate estimation. The goodness-of-fit ($-\ln L$) values for the optimal decision-bound model, the free base-rate decision-bound model, and the general quadratic classifier are displayed in Table 2. The most parsimonious model for each observer is in bold type, and the percent of responses accounted for by this model is also included. Notice that the percent of responses accounted for is quite high, averaging 89% across observers. Since observers were sensitive to base-rate information, the next question to ask is whether each observer's knowledge of base-rates, along with his/her knowledge of the category distribution parameters, was optimal. If so, the optimal decision-bound model should provide the most parsimonious account of the data. Across all base-rate conditions, approximately 25% of the observers did use the optimal decision bound (albeit in the presence of perceptual and criterial noise). Restricting attention to the unequal base-rate conditions, approximately 21% (i.e., 5) of the observers learned the optimal decision bound (4 from the moderate base-rate conditions, and 1 from the large base-rate conditions). Thus, even when the optimal decision bound was highly nonlinear, a large number of observers gained accurate knowledge of the category distributions and base-rates and learned the optimal decision bound.

Suboptimal base-rate estimation. As expected, many observers' data were not best described by the optimal decision-bound model. One possibility that is based on previous research (Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1965; Maddox, 1995; Maddox & Bohil, 1998) is that suboptimal observers accurately extracted information about the category distributions (i.e., used the optimal decision function), but did not use the optimal decision criterion. Some studies have revealed conservative cutoff placement (e.g., Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1965), whereas others have revealed extreme cutoff placement (Maddox, 1995; Maddox & Bohil, 1998). The free base-rate decision-bound model was designed to determine whether observers show conservative or extreme cutoff placement. A second possibility is that suboptimal observers misestimated both the category distribution information and the base-rate information. This hypothesis was tested by the general quadratic classifier. Restricting attention to the unequal base-rate conditions, for 8 of the 19 suboptimal observers, the free base-rate decision-bound model provided the most parsimonious account of the data. The general quadratic classifier provided the most parsimonious account of the remaining 11 observers' data. In summary, nearly half of the suboptimal observers accurately estimated the category distribution information and used a decision bound of the same shape as the optimal bound, but used a suboptimal decision criterion.

Cutoff placement for suboptimal unequal base-rate observers. At this juncture it is important to distinguish between conservative or extreme cutoff placement and under- or overestimation of category base-rates. The former are empirical phenomena that are often observed in data and are quantified using model-based analyses.

The latter are potential psychological implications of the former. In other words, if an observer shows conservative cutoff placement in his/her data, one might argue that this implies an underestimate of the ratio of the category base-rates. This is a reasonable psychological implication of these data, but is certainly not the only implication. Thus, we restrict attention here to discussion of empirical phenomena, and will leave consideration of the possible psychological implications for the Discussion section.

To determine whether observers showed conservative or extreme cutoff placement when the optimal decision function was nonlinear, we examined the $|C_f|$ estimates from the free base-rate decision-bound model for the 19 suboptimal observers.[2] Recall that when Category A is the high base-rate category, C_o is positive, and extreme cutoff placement results when $C_f > C_o$. However, when Category B is the high base-rate category, C_o is negative, and extreme cutoff placement results when $C_f < C_o$. To simplify the presentation, we report the absolute value of C_f and C_o. In line with the results of Maddox (1995; Maddox & Bohil, 1998), the data overwhelmingly support a conclusion of extreme cutoff placement. All 8 suboptimal observers in the moderate base-rate condition and 9 of the 11 suboptimal observers in the large base-rate condition evidenced extreme cutoff placement. In the moderate base-rate condition, $|C_o| = 0.693$, and in the large base-rate condition, $|C_o| = 1.099$. The $|C_f|$ estimates (averaged across suboptimal observers) were 2.415 and 3.686 in the moderate and large base-rate conditions, respectively. In other words, (on average) the suboptimal observers used a decision criterion that was approximately 3.5 times larger than the optimal decision criterion. This, along with similar findings by Maddox (1995; Maddox & Bohil, 1998) with linear optimal decision bounds, suggests that extreme cutoff placement is a robust phenomenon when category base-rates are unequal, stimuli are multidimensional, and observers are highly experienced.

Figure 5 displays plots of the responses for a representative observer from each of the unequal base-rate conditions, along with the optimal decision bound (solid curve) and best fitting free base-rate decision bound (broken curve). Each of these observers showed extreme cutoff placement. The $|C_o|$ and $|C_f|$ values for each of these observers are as follows: 2:1 base-rate Observer C/1, $|C_o| = 0.693$ and $|C_f| = 1.922$ (Figure 5a); 1:2 base-rate Observer C/2, $|C_o| = 0.693$ and $|C_f| = 3.111$ (Figure 5b); 3:1 base-rate Observer R/1, $|C_o| = 1.099$ and $|C_f| = 2.148$ (Figure 5c); and 1:3 base-rate Observer R/1, $|C_o| = 1.099$ and $|C_f| = 3.535$ (Figure 5d).

The $|C_f|$ analyses are based on estimates from a version of decision-bound theory. Thus, their validity is intimately tied to the validity of decision-bound theory. Although the excellent fits of the data suggest that decision-bound theory provides a reasonable description of performance in this task and support a conclusion of extreme cutoff placement in the unequal base-rate conditions, it would be advantageous to use a model-free estimate of the observer's cutoff in order to bolster our claim that extreme cutoff place-
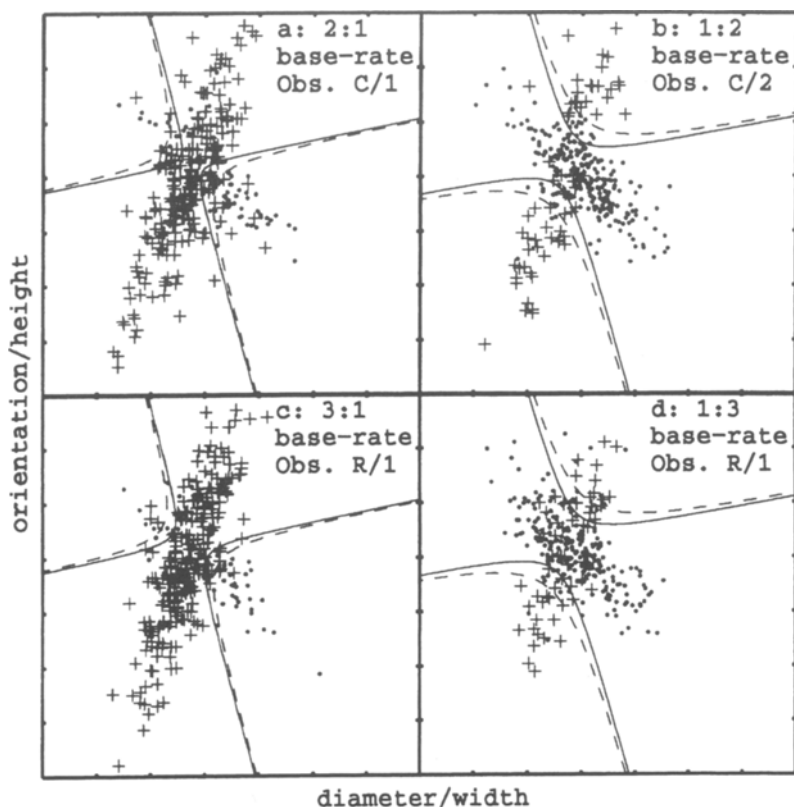
Figure 5. Response data for a representative observer from the (a) 2:1 base-rate condition, Observer C/1; (b) 1:2 base-rate condition, Observer C/2; (c) 3:1 base-rate condition, Observer R/1; and (d) 1:3 base-rate condition, Observer R/1. In each plot, the plus signs denote Category A responses, and the points denote Category B responses. In addition, the solid curve represents the optimal decision bound, and the broken curve represents the best fitting free base-rate decision bound.

ment exists in the unequal base-rate conditions. We develop a "model-free" estimate next.

Observed probability of responding with the "high base-rate" category as a function of distance to the optimal bound. Maddox (1995; Maddox & Bohil, 1998) developed a procedure that estimated the probability of a "high base-rate" category response for different regions of the stimulus space. Basically, Maddox divided the stimulus space into several nonoverlapping regions, all parallel to the optimal decision bound, but at different distances from the optimal bound. Maddox then estimated the probability of the high base-rate category response in each region in an attempt to determine where the probability changed from a value greater than .5 to a value less than .5. If the observer is using the optimal decision bound, the probability of the high base-rate category response will be greater than .5 on one side of the optimal bound and will be less than .5 on the other side of the optimal bound. Therefore, the probabilities will change from values greater than .5 to values less than .5 as one crosses the optimal decision bound. However, suppose the observer shows extreme cutoff placement. In this case, the observer is not using the optimal decision bound, but rather a decision bound that

yields more high base-rate responses. In this case, the probability of the high base-rate category response will remain greater than .5 as one crosses the optimal decision bound. In both studies, these analyses confirmed the $|C_f|$ conclusion of extreme cutoff placement (Maddox, 1995; Maddox & Bohil, 1998).

A similar procedure was used in the present study. Specifically, we divided the stimulus space into several nonoverlapping regions; each region was bounded by two quadratic functions, each of the same shape as the optimal decision bound, but with a different constant term. Region 1 had as its lower bound the optimal quadratic function and as its upper bound a quadratic function of the same shape as the optimal, but with a constant term that differed by 0.4 units in the direction indicative of extreme cutoff placement. Region 2 had as its lower bound the upper bound from Region 1 and an upper bound indicative of even more extreme cutoff placement. All other regions were defined in a similar manner. Notice that the lower bound for Region 1 is the optimal decision bound. Thus, if the observer is using the optimal decision bound, the probability of a high base-rate category response should be less than .5 in Region 1 and should decrease monotonically for Re-

gions 2–10. On the other hand, if the observer shows extreme cutoff placement, the probability of a high base-rate category response should be greater than .5 for Region 1 and should decrease monotonically for the other regions. The data support a hypothesis of extreme cutoff placement. In the moderate and large base-rate conditions, the probability of a high base-rate category response was greater than .5 in Region 1 (moderate condition, .545; large condition, .523) and fell off monotonically for the other regions.

To summarize, it appears that extreme cutoff placement is a robust phenomenon when category base-rates are unequal, stimuli are multidimensional, observers are experienced, and single observer analyses are performed. Extreme cutoff placement was found (1) for several different linear and nonlinear optimal decision bounds, (2) across a wide variety of stimulus dimensions (rectangles varying in height and width; circles varying in size and orientation of an embedded radial line; a line varying in length and orientation; and a dot varying in horizontal and vertical location), and (3) for simple and complex category structures.

Three model-based instantiations of Kruschke's (1996) theory. Three additional hypotheses were tested, all of which were motivated by Kruschke's work. As noted, Kruschke hypothesized that (1) base-rate information is learned and is applied consistently, (2) the high base-rate category is learned before the low base-rate category, and (3) the high base-rate category is learned by its typical features, while the low base-rate category is learned by its distinctive features. Because Kruschke's theory was developed for discrete-valued stimuli, and our stimuli are continuous-valued, no direct test was possible. The approach taken here was to develop some reasonable continuous-valued analogues of Kruschke's theory that could serve as a preliminary test of his theory in the domain of continuous-valued dimensions. All three hypotheses we tested assumed that the high base-rate category was learned before the low base-rate category. Since the categories were normally distributed, this required the observer to learn something about the central tendency and exemplar variability within a category. Our second and third hypotheses were most similar in spirit to those of Kruschke because in addition they attempted to incorporate a continuous-valued analogue of Kruschke's typical versus distinctive feature dichotomy.

We will use Figure 6 as a guide in developing the various model-based instantiations of Kruschke's (1996) theory. The six panels in the left-hand column of Figure 6 illustrate the assumed stimulus regions for each model. In line with Figure 3, we will use the labels A and B to denote the true category distributions. In some of the models, we assume that the observer does not use an accurate estimate of the true category distribution, but rather makes some incorrect assumption about the category. These will be called "inferred" categories, and will be denoted by the labels A′ and B′. The six panels in the right-hand column of Figure 6 illustrate the assumed decision bounds and re-

sponse regions. Again, in line with Figure 3 we will use the labels A and B to denote the response regions.

Hypothesis 1 postulates that the observer accurately estimates the category means and knows that Category A is displaced to the left of Category B. However, the observer incorrectly assumes that the low base-rate category has the same covariance structure as the high base-rate category (i.e., is characterized by a contour of equal likelihood with the same shape and orientation). This hypothesis is depicted visually in Figures 6a and 6c. In Figure 6a, Category A is the high base-rate category. We label the Category A contour with an A to denote a correctly assumed covariance structure, whereas we label the Category B contour with a B′ to denote an incorrectly inferred covariance structure. Figure 6c depicts the same hypothesis for the case in which Category B is the high base-rate category. When the stimulus regions are as shown in Figures 6a and 6c, the optimal decision bound would be linear. The decision bound and response regions associated with Figures 6a and 6c are depicted in Figures 6b and 6d, respectively. To test this hypothesis, we developed a linear decision-bound model whose slope and intercept were estimated from the data. The only constraint was that the slope be positive when Category A was the high base-rate category (see Figure 6b) and that the slope be negative when Category B was the high base-rate category (see Figure 6d). Thus, the model did not require that the observer have a perfect estimate of the high base-rate category distribution. The model required only that the observer have a reasonable estimate. We called this the *single linear bound* model.

Hypothesis 2 assumes that the observer learns the high base-rate category and effectively partitions the perceptual space into two regions: one "structured" region associated with the high base-rate category and one "unstructured" region associated with the low base-rate category. The stimulus regions for this case are depicted in Figures 6e and 6g, and the associated decision bounds and response regions are depicted in Figures 6f and 6h, respectively. In Figure 6e, Category A is the high base-rate category and is assumed to be accurately estimated (hence the label A). Category B is assumed to be "unstructured," which is incorrect (hence the label B′). To instantiate the notion of a "structured" high base-rate response region and an "unstructured" low base-rate response region, we assumed that the observer constructed an elliptical decision bound similar in shape to the contour of equal likelihood (see Figure 6f). An instantiation of this model for the case in which Category B is the high base-rate category is depicted in Figures 6g and 6h. To test this hypothesis, we developed a quadratic decision-bound model whose parameters were estimated from the data. The only constraint was that the decision bound be elliptical and have a positive tilt when Category A was the high base-rate category (see Figure 6f) and have a negative tilt when Category B was the high base-rate category (see Figure 6h). Thus, as in Hypothesis 1, the model does not require that
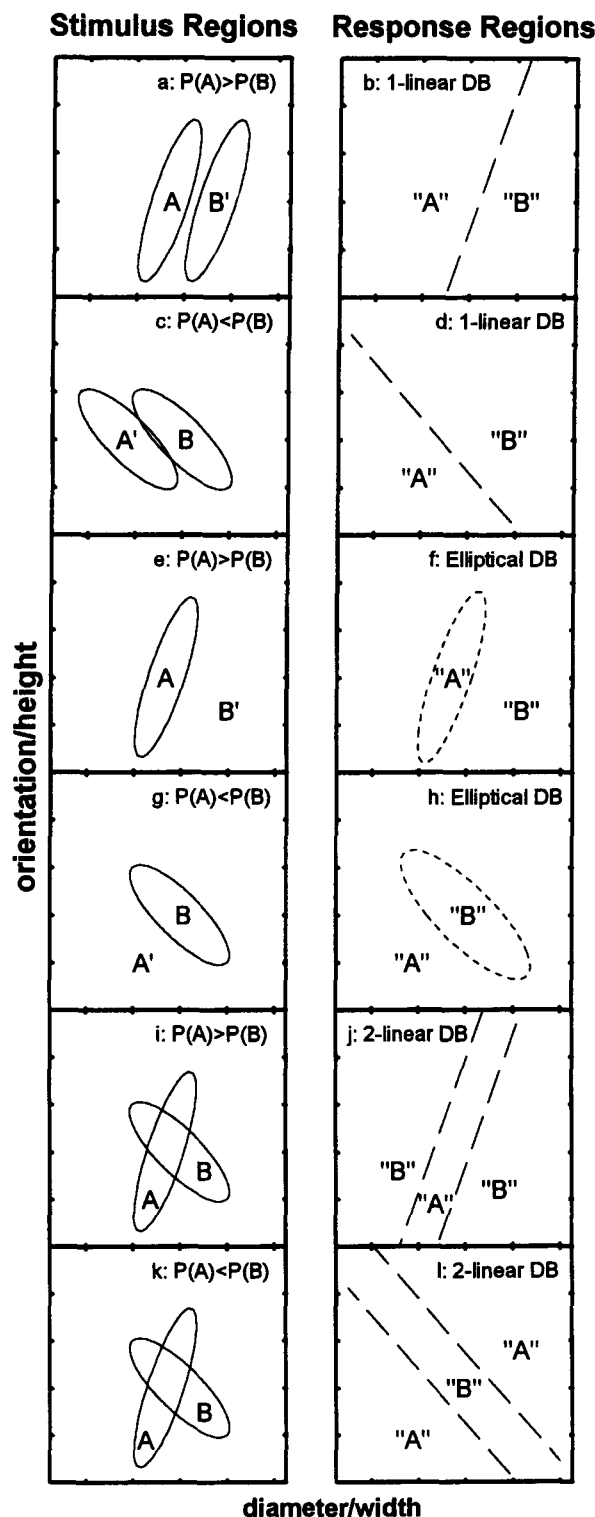
Stimulus Regions   Response Regions

into a structured high base-rate category response region and an unstructured low base-rate category response region, the observer is assumed to partition the space into three response regions: one assigned to the high base-rate category and the remaining two assigned to the low base-rate category. The assumptions about the stimulus regions are depicted in Figures 6i and 6k. In short, the observer is assumed to have accurate knowledge of both category structures (hence the labels A and B). The decision bounds and response regions for the high A and high B base-rate conditions are depicted in Figures 6j and 6l, respectively. Essentially, the observer postulates that the low base-rate category is actually composed of two subcategories that are distinct from each other. This seems reasonable given the category structures used in our study. To implement this model, we assumed that the partitions between the high base-rate and two low base-rate subcategories were linear. This hypothesis is depicted in Figure 6j (Category A high base-rate) and Figure 6l (Category B high base-rate). We called this the *bilinear decision-bound* model.

The single linear, bilinear, and elliptical decision-bound models were applied to the final session data from each of the 24 observers in an unequal base-rate condition. Not surprisingly, the single linear decision-bound model provided a poor description of the data. For all 24 observers the bilinear and elliptical decision-bound models provided significantly better accounts of the data. Both the elliptical and bilinear decision-bound models provided a good

the observer have a perfect estimate of the high base-rate category distribution; the model requires only that the observer have a reasonable estimate. We called this the *elliptical decision-bound* model.

Hypothesis 3 was similar to Hypothesis 2; however, instead of assuming that the observer partitions the space

Table 3
Goodness-of-Fit (−lnL) Values for First and
Last Experimental Sessions Averaged Across Observers

Condition	OPT (2 Parameters)		FRB (3 Parameters)		GQC (7 Parameters)	
	First	Last	First	Last	First	Last
1:1 Base-rate	155.30	102.37	151.84	98.97	124.00	93.11
1:2 Base-rate	171.83	93.39	144.34	88.23	123.99	84.71
2:1 Base-rate	149.19	107.41	127.94	91.70	116.80	83.78
Moderate mean	160.51	100.40	136.14	89.97	120.40	84.25
1:3 Base-rate	148.21	95.77	114.75	72.34	99.00	66.23
3:1 Base-rate	148.72	63.36	139.17	58.95	118.17	52.69
Large mean	148.47	79.57	126.96	65.65	108.59	59.46

Note—OPT, optimal decision-bound model; FRB, free base-rate optimal decision-bound model; GQC, general quadratic classifier. Moderate mean refers to the moderate base-rate ratios (1:2 and 2:1). Large mean refers to the large base-rate ratios (1:3 and 3:1). The final session was the third session for 2 observers and the fourth session for the other 30 observers.

account of the data. Averaged across the 24 observers, both models accounted for 87% of the responses. This compares quite favorably with 90% of the responses accounted for by the most parsimonious decision-bound model (averaged across the 24 observers). Even so, on the basis of the maximum likelihood goodness-of-fit measure, the most parsimonious decision-bound model provided a better account of the data than either the elliptical or bilinear models for 23 of the 24 observers.

It is important to note that this is not meant as a complete and rigorous test of Kruschke's (1996) theory in the domain of continuous-valued dimensions. We have developed only a few reasonable instantiations of Kruschke's theory and do not feel that these data provide convincing evidence for or against Kruschke's theory. In support of decision-bound theory, the decision-bound models outlined in Figure 1 provide the most accurate account of the data, outperforming all three proposed instantiations of Kruschke's theory. Even so, the success of the elliptical and bilinear models suggests that the notion that high and low base-rate categories are learned in a different manner might have some validity even in the domain of continuous-valued dimensions.

Changes in Performance Over Sessions

A number of results suggest that performance became more optimal as observers gained experience with the categories. However, the data suggest also that performance reached asymptote at a suboptimal level. For example, the accuracy rates showed a general increase from the first to the last experimental session in each condition (see Figure 7a) but reached asymptote at a level well below that predicted by the optimal classifier. To further examine the changes in performance with experience, the three decision-bound models depicted in Figure 1 were applied to all four experimental sessions for each observer. A comparison of the first and final session fits (averaged across observers) are reported in Table 3. The primary results are as follows. First, the fits of the general quadratic classifier, the free base-rate decision-bound model, and the optimal

decision-bound model were more similar when applied to the final session data (average difference across models = 10.91 −lnL units) than when applied to the first session data (average difference across models = 25.45 −lnL units), suggesting that observers' estimates of the base-rate difference improved across sessions. Second, the optimal decision-bound model evidenced the largest improvement in fit from the first to the final experimental session, further supporting the hypothesis that responding becomes more optimal with experience.

Changes in decision criterion. To examine the time course of extreme cutoff placement, the C_f estimates from the free base-rate decision-bound model were examined for each of the four sessions. Figure 7b plots the ratio of the C_f estimates with the optimal decision criterion C_o over sessions for the unequal base-rate conditions. A C_f/C_o ratio of greater than one indicates extreme cutoff placement, and a ratio of less than one indicates conservative cutoff placement. When $|C_f/C_o| = 1$, the optimal decision criterion is being used. The C_f/C_o ratio was largest for the first experimental session and decreased during the second (and often third) session, then reached asymptote at a level still indicative of extreme cutoff placement.

DISCUSSION

The present report examined the effects of unequal category base-rates and nonlinear optimal decision bounds on the optimality of human categorization performance. Five base-rate ratios were combined factorially with two qualitatively different types of stimulus dimensions. Each observer completed several experimental sessions to ensure adequate levels of experience, and all analyses were conducted at the level of the individual observer. The analyses centered around the application of a series of nested decision-bound models that were derived from decision-bound theory. Decision-bound theory postulates that observers attempt to respond optimally, but are generally suboptimal due to several inherent weaknesses in perceptual and cognitive processing.

a: Accuracy Rate Across Sessions

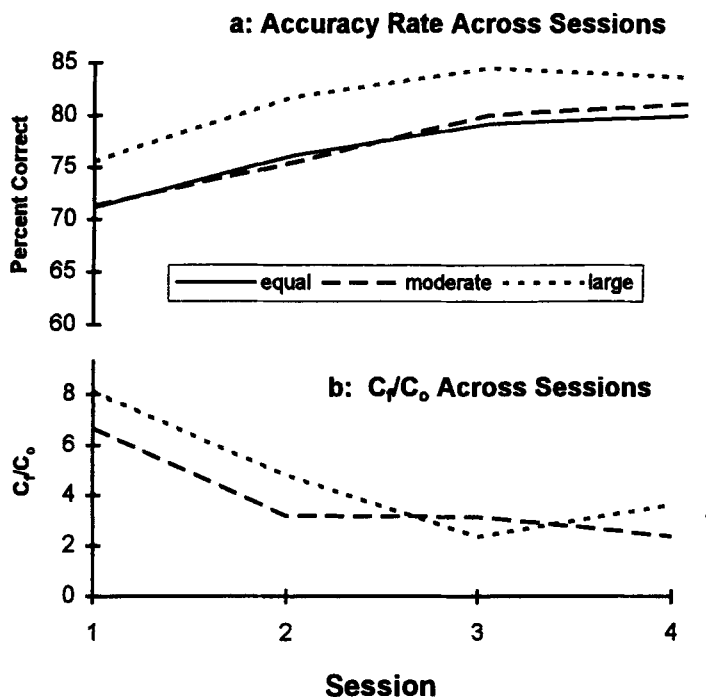## b: C_f/C_o Across Sessions

Session

Figure 7. (a) Accuracy rate and (b) the C_f/C_o ratio as a function of experimental session. The points for Session 4 are based on the data for 30 observers (2 observers did not complete the fourth session). The points for Sessions 1, 2, and 3 are based on the data for all 32 observers.

All observers demonstrated knowledge of the category base-rate manipulations, even during their first experimental session. In fact, by their final experimental session, approximately one fourth of the observers responded in accordance with the actual base-rate ratios, as evidenced by their use of the optimal decision function and the optimal decision criterion. Although the nonlinear optimal decision bound differed across conditions and two qualitatively different types of stimulus dimensions were utilized, most observers showed the same pattern of responding—namely, extreme cutoff placement. These findings are supported by extensive model-based analyses, as well as by "model-free" indices of performance.

These findings converge with those from several conditions in which the optimal decision bound was linear (Maddox, 1995; Maddox & Bohil, 1998). Across all three studies, the results seem clear: When faced with a multidimensional perceptual categorization problem in which the category base-rates are unequal, observers (given enough exposure to the categories) are quite adept at estimating the category distribution parameters that determine the shape of the optimal decision bound (whether linear or nonlinear). In addition, observers are highly sensitive to base-rate manipulations, but tend to use a decision criterion that is more extreme than that predicted by the optimal classifier.

Psychological Implications
of Extreme Cutoff Placement

One of the more interesting aspects of the present analyses was the continued support for the prevalence of

extreme cutoff placement in multidimensional perceptual categorization tasks that incorporate unequal category base-rates. Extreme cutoff placement is an empirical phenomenon that was observed in our data. In this section, we examine several possible psychological implications of extreme cutoff placement. One reasonable implication is that observers overestimate the true base-rate difference and use a decision criterion that is more extreme than the optimal decision criterion. Although this hypothesis cannot be rejected by the data, we know of no studies that directly examine the accuracy of observers' base-rate estimation that could corroborate this finding. In addition, the data from several observers was fit best by the optimal decision-bound model. It is likely that these observers accurately estimated the base-rate ratio, which would seem at odds with the hypothesis that observers tend to overestimate the true base-rate difference. Even so, more work is needed to determine whether observers actually overestimate base-rate ratios.

A second possibility is that observers are good at estimating base-rate ratios (at least after extensive exposure to the categories), but often show a bias toward the high base-rate category. Our speculation is that this bias is most likely to be invoked when the observer is unsure about the category membership of an exemplar. Recall that observers are assumed to compare the output of their decision function with their decision criterion. If the output of the decision function is greater than the decision criterion, they give one response, and if it is less than the decision criterion, they give another response (see Equation 2). The

observer will be unsure about category membership when the output of the decision function is similar to that of the decision criterion. In these cases, the observer might be biased to give the high base-rate response. In fact, when the base-rates are unequal this is a reasonable strategy for an observer faced with uncertainty because more often than not the high base-rate response will be the correct response. Call this an *uncertainty-driven* response bias. Unfortunately, it is difficult to explicitly model a response bias in this study because it would be nonidentifiable with the estimated decision criterion. Even so, several predictions can be derived from this uncertainty-driven response bias hypothesis that have some support in the data.

First, we would expect to find extreme cutoff placement in the estimated decision criterion. If the observer is more likely to give the high base-rate response when the output of his/her decision function is similar to his/her decision criterion, then at the level of the data it will appear as if the observer is using a more extreme decision criterion. Second, we would expect the observer's level of experience with the categories to affect the magnitude of extreme cutoff placement. An observer with little exposure to the categories will be faced with many exemplars whose category membership is unclear (i.e., will be in a state of high uncertainty). If the observer is assumed to give the high base-rate response under these conditions, we would expect to find a large proportion of uncertainty-driven high base-rate responses early in learning. As the observer gains more experience with the categories, uncertainty should decrease and the proportion of uncertainty-driven high base-rate responses should decrease. Thus, the magnitude of extreme cutoff placement should be larger for the novice observer and should decrease as the observer gains experience with the categories. Both of these predictions are supported by the data.

A third prediction from this hypothesis is that some observers might learn to suppress the uncertainty-driven response bias as they gain experience with the categories. Observers using the same strategy as the optimal classifier will never encounter a situation in which they are uncertain about the response that should be given and thus should never show an uncertainty-driven response bias. At the level of the data, then, we would expect those observers who ultimately learned the optimal decision bound to initially show extreme cutoff placement (perhaps due to an uncertainty-driven response bias early in learning). The magnitude of extreme cutoff placement should decrease as these observers learn to suppress the response bias, until extreme cutoff placement disappears and the data are best described by the optimal decision-bound model. To test this hypothesis we examined the C_f/C_o estimates for the 5 observers from the unequal base-rate conditions whose final session data were fit best by the optimal decision-bound model. If the uncertainty-driven response bias hypothesis is correct, then the C_f/C_o estimates should be greater than one during the first session, suggesting that a response bias exists, and should decrease over sessions to a value near one, indicative of no response bias. For all 5

observers, the C_f/C_o estimates were larger than one during the first experimental session and decreased over sessions. Averaged across the 5 observers, the C_f/C_o estimates decreased monotonically across sessions, yielding values of 11.48, 2.82, 1.61, and 1.15 for Sessions 1–4, respectively. Thus, these analyses provide support for the hypothesis that the prevalence of extreme cutoff placement might be due to an uncertainty-driven response bias whose magnitude decreases as the observer gains experience with the categories. Although this response bias hypothesis provides an interesting account of the data, clearly other explanations are possible, and more work is needed before a definitive conclusion is reached. A major focus of future research should be to develop rigorous quantitative models of perceptual categorization performance that can account more precisely for the prevalence of extreme cutoff placement.

Another possibility is that extreme cutoff placement results because the high and low base-rate categories are learned in different ways. As suggested by Kruschke (1996), the high base-rate category may be learned first, and then the distinctive aspects of the low base-rate category might be used to distinguish it from the high base-rate category. Three model-based instantiations of this hypothesis were tested and two of the three received some support. Even so, in most cases a decision-bound model that assumes that the observer attempts to respond optimally but cannot because of several perceptual and cognitive limitations was superior.

Conservative Versus Extreme Cutoff Placement

The present study, along with studies conducted by Maddox (1995; see also Maddox & Bohil, 1998), suggests that extreme cutoff placement is likely to result when category base-rates are unequal. However, a series of previous studies (Healy & Kubovy, 1981; Lee & Janke, 1965; Ulehla, 1966) suggests that conservative cutoff placement is likely to result when category base-rates are unequal. Why is extreme cutoff placement so prevalent in one set of studies and conservative cutoff placement so prevalent in another set of studies? Maddox outlined several differences between the two sets of studies that might account for the discrepant findings. In many cases, the previous studies (Healy & Kubovy, 1981; Lee & Janke, 1965; Ulehla, 1966) used fairly inexperienced observers, analyzed the data collapsed across observers, and used unidimensional stimuli. Maddox (1995; Maddox & Bohil, 1998) and the present study used highly experienced observers, analyzed the data at the individual observer level, and used two-dimensional stimuli. Thus, one possibility is that conservative cutoff placement is more likely to occur when stimuli are unidimensional and the data are averaged across novice observers, whereas extreme cutoff placement is more likely to occur when the stimuli are multidimensional, observers are highly experienced, and single observer analyses are performed. Several aspects of the current data argue against this hypothesis. First, if novices show conservative cutoff placement, we would expect to find conservative cutoff placement early in learning and

extreme cutoff placement later in learning. This pattern was not found. Second, if averaging data causes conservative cutoff placement, we would expect to find a mixture of conservative and extreme cutoff placement in the single observer data. Again this was not found.

Other potentially important differences between the studies conducted in our laboratory and previous studies have been identified. For example, the distinction between unidimensional and multidimensional stimuli might be important. In addition, the types of stimulus dimension and d' type measures differ across studies. We are currently examining each of these variables in a controlled setting, the goal being to identify the variable (or variables) that predicts cases in which conservative or extreme cutoff placement will result.

In conclusion, the present study extends our understanding of the use of base-rate information in multidimensional perceptual categorization to several situations in which the categories are complex and the optimal decision bound is highly nonlinear. Because greater category complexity leads to highly nonlinear decision functions, decision-bound theory predicts that learning of the decision function should be slowed by category complexity. On the other hand, category base-rates affect only the decision criterion and are independent of category complexity. In line with these predictions, learning of nonlinear decision functions was more difficult and time-consuming in this study as compared with the linear decision functions in Maddox (1995; Maddox & Bohil, 1998). Even so, as predicted, base-rate learning was similar in the two cases. Specifically, in line with the results using simple category structures (Maddox, 1995; Maddox & Bohil, 1996), most observers showed extreme cutoff placement.

REFERENCES

ASHBY, F. G. (1992a). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449-484). Hillsdale, NJ: Erlbaum.

ASHBY, F. G. (1992b). Multivariate probability distributions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 1-34). Hillsdale, NJ: Erlbaum.

ASHBY, F. G., & GOTT, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14, 33-53.

ASHBY, F. G., & LEE, W. W. (1991). Predicting similarity and categorization from identification. *Journal of Experimental Psychology. General*, 120, 150-172.

ASHBY, F. G., & LEE, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. In S. C. Masin (Ed.), *Foundations of perceptual theory* (pp. 369-397). New York: Elsevier, North-Holland.

ASHBY, F. G., & MADDOX, W. T. (1990). Integrating information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 598-612.

ASHBY, F. G., & MADDOX, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception & Performance*, 18, 50-71.

ASHBY, F. G., & MADDOX, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, 37, 372-400.

ASHBY, F. G., & MADDOX, W. T. (1994). A response time theory of perceptual separability and perceptual integrality in speeded classification. *Journal of Mathematical Psychology*, 33, 423-466.

ASHBY, F. G., & MADDOX, W. T. (1998). Stimulus categorization. In

M. H. Birnbaum (Ed.), *Measurement, judgment, & decision making* (pp. 251-301). New York: Academic Press.

ASHBY, F. G., MADDOX, W. T., & LEE, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science*, 5, 144-150.

ASHBY, F. G., & PERRIN, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review*, 95, 124-150.

ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, 93, 154-179.

BALLA, J. I. (1982). The use of critical cues and prior probability in decision-making. *Methods of Information in Medicine*, 21, 9-14.

BALLA, J. I., ELSTEIN, A., & GATES, P. (1983). Effects of prevalence and test diagnosticity upon clinical judgments of probability. *Methods of Information in Medicine*, 22, 25-28.

CASSCELLS, B. S., SCHOENBERG, A., & GRABOYS, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine*, 229, 999-1000.

EDGELL, S. E., CASTELLAN, JR., N. J., ROE, R. M., BARNES, J. M., NG, P. C., BRIGHT, R. D., & FORD, L. A. (1996). Irrelevant information in probabilistic categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 1463-1481.

EDGELL, S. E., & HENNESSEY, J. E. (1980). Irrelevant information and utilization of event base rates in nonmetric multiple-cue probability learning. *Organizational Behavior & Human Performance*, 26, 1-6.

ESTES, W. K. (1950). The problem of inference from curves based on group data. *Psychological Bulletin*, 53, 134-140.

ESTES, W. K., CAMPBELL, J. A., HATSOPOULOS, N., & HURWITZ, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 556-571.

FELFOLDY, A. (1974). Repetition effects in choice reaction time to multidimensional stimuli. *Perception & Psychophysics*, 15, 453-459.

FELFOLDY, G., & GARNER, W. R. (1971). The effects on speeded classification of implicit and explicit instructions regarding redundant dimensions. *Perception & Psychophysics*, 9, 289-292.

FRIEDMAN, D., MASSARO, D. W., KITZIS, S. N., & COHEN, M. M. (1995). A comparison of learning models. *Journal of Mathematical Psychology*, 39, 164-178.

GARNER, W. R., & FELFOLDY, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1, 225-241.

GEISLER, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, 96, 267-341.

GIGERENZER, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, 103, 592-596.

GIGERENZER, G., HELL, W., & BLANK, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception & Performance*, 14, 513-525.

GIGERENZER, G., & HOFFRAGE, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704.

GIGERENZER, G., & MURRAY, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.

GLUCK, M. A., & BOWER, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 117, 225-244.

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

HEALY, A. F., & KUBOVY, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning & Memory*, 7, 344-354.

HOLYOAK, K. J., & SPELLMAN, B. A. (1993). Thinking. In L. W. Porter & M. R. Rosenzweig (Eds.), *Annual review of psychology* (Vol. 44, pp. 265-315). San Diego: Academic Press.

KAHNEMAN, D. K., & TVERSKY, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582-591.

KAHNEMAN, D. K., & TVERSKY, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

KRUSCHKE, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 3-26.

KUBOVY, M., & HEALY, A. F. (1977). The decision rule in probabilistic

categorization: What it is and how it is learned. *Journal of Experimental Psychology: General*, 106, 427-466.

LEE, W., & JANKE, M. (1964). Categorizing three externally distributed stimulus samples for three continua. *Journal of Experimental Psychology*, 68, 376-382.

LEE, W., & JANKE, M. (1965). Categorizing externally distributed stimulus samples for unequal molar probabilities. *Psychological Reports*, 17, 79-90.

LEE, W., & ZENTALL, T. R. (1966). Factorial effects in the categorization of externally distributed stimulus samples. *Perception & Psychophysics*, 1, 120-124.

MADDOX, W. T. (1995). Base-rate effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21, 288-301.

MADDOX, W. T., & ASHBY, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, 53, 49-70.

MADDOX, W. T., & ASHBY, F. G. (1996). Perceptual separability, decisional separability, and the identification–speeded classification relationship. *Journal of Experimental Psychology: Human Perception & Performance*, 22, 795-817.

MADDOX, W. T., & ASHBY, F. G. (1998). Selective attention and the formation of linear decision boundaries: Comment on McKinley and Nosofsky (1996). *Journal of Experimental Psychology: Human Perception & Performance*, 24, 301-321.

MADDOX, W. T., & BOHIL, C. J. (1998). *Base-rate and payoff effects in multidimensional perceptual categorization*. Manuscript submitted for publication.

MEDIN, D. L., & EDELSON, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General*, 117, 68-85.

MORRISON, D. F. (1967). *Multivariate statistical methods*. New York: McGraw-Hill.

NOSOFSKY, R. M. (1985). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics*, 38, 415-432.

NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-classification relationship. *Journal of Experimental Psychology: General*, 115, 39-57.

NOSOFSKY, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13, 87-109.

NOSOFSKY, R. M., CLARK, S. E., & SHIN, H. J. (1989). Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 282-304.

NOSOFSKY, R. M., KRUSCHKE, J. K., & McKINLEY, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18, 211-233.

SCHONEMANN, P. H. (1977). Similarity of rectangles. *Journal of Mathematical Psychology*, 16, 161-165.

SHEPARD, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54-87.

SPELLMAN, B. A. (1993). Implicit learning of base rates. *Psycoloquy* [On-line serial], 4(61). Available: gopher: //gopher.Princeton.EDU: 70/0ftp%3Aprinceton.edu@/pub/harnad/Psycoloquy/1993.volume.4/psyc.93.4.61.base-rate.4.spellman

STEVENS, S. S. (1961). Psychophysics of sensory function. *American Scientist*, 48, 226-252.

THOMAS, R. D. (1995). Gaussian general recognition theory and perceptual independence. *Psychological Review*, 102, 192-200.

TVERSKY, A., & KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

ULEHLA, Z. J. (1966). Optimality of perceptual decision criteria. *Journal of Experimental Psychology*, 71, 564-569.

WALLSTEN, T. S. (1981). Physician and medical student bias in evaluating diagnostic information. *Medical Decision Making*, 1, 145-164.

WEBER, E. U., BOCKENHOLT, U., HILTON, D. J., & WALLACE, B. (1993). Determinants of diagnostic hypothesis generation: Effects of information, base-rate, and experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 19, 1-14.

WICKENS, T. D. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco: Freeman.

WYSZECKI, G., & STILES, W. S. (1967). *Color science: Concepts and methods, quantitative data and formulas*. New York: Wiley.

NOTES

1. Perceptual noise, in the visual case, is caused by fluctuations in the number of photons that reach the cornea and spontaneous activity within the central nervous system. For example, it is well known that the number of photons emitted by a light source of constant intensity and constant duration is well described by a Poisson distribution (Geisler, 1989; Wyszecki & Stiles, 1967). Because the mean and variance of the Poisson distribution are equal, the variability in the number of photons emitted increases with stimulus intensity.

2. The most rigorous test of conservative and extreme cutoff placement would require the free base-rate decision bound model to provide the most parsimonious account of the data. However, only 8 of the 19 suboptimal observers' data were most parsimoniously accounted for by this model, and the remaining 11 were most parsimoniously accounted for by the general quadratic classifier. Even so, we decided to include all 19 observers in our subsequent analyses for the following reasons. First, for the 11 general quadratic classifier observers, the free base-rate model and the general quadratic classifier fits were similar (average fit: free base-rate model = 79.71, general quadratic classifier = 69.08). Second, the percent of responses accounted for by each model was high, and the percents were very similar (average percent of responses accounted for: free base-rate model = 89%, general quadratic classifier = 90%). Third, the main theoretical conclusions were not affected by including these 11 observers in the subsequent analyses.