JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2002, **78,** 567–595 NUMBER 3 (NOVEMBER)

TOWARD A UNIFIED THEORY OF DECISION CRITERION LEARNING IN PERCEPTUAL CATEGORIZATION

W. TODD MADDOX

UNIVERSITY OF TEXAS, AUSTIN

Optimal decision criterion placement maximizes expected reward and requires sensitivity to the category base rates (prior probabilities) and payoffs (costs and benefits of incorrect and correct responding). When base rates are unequal, human decision criterion is nearly optimal, but when payoffs are unequal, suboptimal decision criterion placement is observed, even when the optimal decision criterion is identical in both cases. A series of studies are reviewed that examine the generality of this finding, and a unified theory of decision criterion learning is described (Maddox & Dodd, 2001). The theory assumes that two critical mechanisms operate in decision criterion learning. One mechanism involves competition between reward and accuracy maximization: The observer attempts to maximize reward, as instructed, but also places some importance on accuracy maximization. The second mechanism involves a flat-maxima hypothesis that assumes that the observer's estimate of the reward-maximizing decision criterion is determined from the steepness of the objective reward function that relates expected reward to decision criterion placement. Experiments used to develop and test the theory require each observer to complete a large number of trials and to participate in all conditions of the experiment. This provides maximal control over the reinforcement history of the observer and allows a focus on individual behavioral profiles. The theory is applied to decision criterion learning problems that examine category discriminability, payoff matrix multiplication and addition effects, the optimal classifier's independence assumption, and different types of trial-by-trial feedback. In every case the theory provides a good account of the data, and, most important, provides useful insights into the psychological processes involved in decision criterion learning.

Key words: decision criterion learning, category learning, base rates, cost–benefits, optimality, multivariate signal-detection theory

---

Categorization is a primary component of the behavior of all organisms. Rats categorize bits of food as ''large'' or ''small,'' with small pieces being eaten immediately and large pieces being hoarded (Wishaw, 1990; Wishaw & Tomie, 1989). The feeding deer must categorize a sound, like leaves rustling, as indicative of ''a hunter approaching'' or ''no hunter approaching,'' with an approaching hunter leading the deer to cease feeding and to run. The expedition doctor camped out near the summit of K2 must categorize a climber's difficulty breathing as a sign of ''pulmonary edema'' or ''exhaustion,'' with a pulmonary edema diagnosis leading to immediate retreat. These are all categorization problems because in every case there are many (generally an infinite number of) information states, but only a few (often two) courses of action. Categorization performance is governed by the organism's experience with the environment and the reinforcing consequences of the decisions that they make. With experience, organisms become adept at many categorization tasks. If they did not, they would die. In light of this fact, it is reasonable to hypothesize that in many domains, human and other organisms' categorization performance is nearly optimal (Ashby & Maddox, 1998). Although *optimality* can be defined in many ways, a common definition is performance that maximizes expected reward (Green & Swets, 1967). To maximize expected reward, one needs explicit knowledge of the distributional properties of the categories, or essentially an infinite amount of experience. Neither is likely. Even so, fairly small samples yield reasonable estimates of the relevant information. Of course, as events

---

This research was supported in part by National Science Foundation Grant SBR-9796206 and National Institute of Mental Health Grant 5 R01 MH59196. I thank Greg Ashby, Jerry Busemeyer, Ido Erev, Neil Macmillan, and Tom Wallsten for many helpful discussions that have greatly influenced my thinking, and Corey Bohil, John Wixted, and two anonymous reviewers for helpful comments on an earlier version of this article.

Correspondence should be addressed to W. Todd Maddox, Department of Psychology, University Station A8000, University of Texas, Austin, Texas 78712 (e-mail: maddox@psy.utexas.edu).

recede into the past, their impact is diminished. Temporal discounting processes (see Mazur, 1988) are therefore relevant, but are not discussed here.

Optimal categorization performance often requires the organism to place a *decision criterion* along the relevant dimension. Using the examples above, the rat must set a criterion on size, the deer must set a criterion on loudness, and the expedition doctor must set a criterion on ease of breathing, with values below the criterion leading to one categorization response, and values above the criterion leading to the other categorization response. The location of the optimal decision criterion is affected by the category base rates (i.e., the prior probability of each category) and the entries in the category payoff matrix (i.e., the costs and benefits associated with correct and incorrect categorization responses). For example, during hunting season the deer's decision criterion might be set at a lower loudness value than during the off-season because hunters are more prevalent during hunting season. Similarly, the decision criterion might be set at a lower loudness value for a mother deer with her fawn than for a deer without offspring because the cost associated with an incorrect "no hunter" decision for the mother deer might be greater than for the deer without offspring.

As outlined in detail below, the location of the optimal decision criterion is determined from the product of the base-rate ratio with the payoff ratio (see Equations 2 and 4 below). For example, suppose that Category A is three times more likely to occur than Category B, the benefit of each correct A response is equal to the benefit of each correct B response, and the cost of each incorrect A response is equal to the cost of each incorrect B response. Under these conditions, the base-rate ratio of Category A to B is 3:1, and the payoff ratio of Category A to B is 1:1 (because the benefits for the two categories are equal and the costs of the two categories are equal). Contrast that with a situation in which Category A and Category B are equally likely to occur, the benefit of each correct A response is three times the benefit of each correct B response, and the cost of each incorrect A response is equal to the cost of each incorrect B response. Under these conditions, the base-rate ratio is 1:1, and the payoff ratio is 3:1 (because the benefit for a correct A response is three times the benefit for a correct B response and the costs of the two categories are equal). Because the product of the base-rate and payoff ratios is 3:1 in both cases, the optimal decision criterion is identical in both cases.

A robust finding in the decision criterion literature is that decision criterion placement in unequal base-rate conditions is consistently closer to optimal than decision criterion placement in unequal payoff conditions, even when the optimal decision criterion is identical across the two conditions (e.g., Green & Swets, 1967; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Ulehla, 1966). Although this empirical result has been replicated many times, the range of experimental contexts in which this phenomenon has been studied is fairly limited. For example, most of this work used a fixed category discriminability, $d' = 1.0$ (where d' is defined as the standardized distance between the category means). In addition, in most of this work, payoff manipulations were instantiated by manipulating the benefit or gain associated with correct responses (usually in the form of points that were converted to money following the experiment), while the cost or loss associated with an incorrect response was fixed at zero. Also, most of these studies examined decision criterion learning when base rates and payoffs were manipulated separately across experimental conditions, while situations in which base rates and payoffs were manipulated simultaneously within the same condition were excluded. Finally, in most of this work the observer received trial-by-trial feedback regarding the correctness of his or her response.

One focus of my research has been to explore the generality of this empirical result by broadening the range of experimental contexts used to examine decision criterion learning with unequal base rates and payoffs. In particular, I have examined decision criterion learning (a) across a wide range of category discriminabilities, (b) in situations in which incorrect responses led to a loss of points, (c) in situations in which base rates and payoffs are manipulated simultaneously within the same experimental context, and

(d) for trial-by-trial feedback based on the objectively correct response versus feedback based on the response made by the optimal classifier. A second focus of my research has been to develop and test specific hypotheses regarding decision criterion learning by comparing human decision criterion learning with that of the optimal classifier (for related work from animal learning, see Alsop, 1998; Herbranson, Fremouw, & Shimp, 1999; Shimp, 1966, 1969, 1973; Shimp, Long, & Fremouw, 1996; see also Staddon, 1992; Staddon & Ettinger, 1989; Stephens & Krebs, 1986). Although in practice performance is often suboptimal, using the optimal classifier as a benchmark has provided many useful insights into the nature of performance suboptimalities and has provided a useful starting point for developing and testing theoretically motivated models of performance. As a result of this work, my colleagues and I have recently proposed, and are now testing, a unified theory of decision criterion learning (Maddox & Dodd, 2001). In fact, predictions derived from the theory helped to guide the choice of experimental contexts outlined above.

This article provides an overview of my recent work on decision criterion learning when base rates and payoffs are manipulated across a wide range of experimental contexts, and reports on the status of a recently proposed unified theory of decision criterion learning. The theory has its roots in signal-detection theory, and thus makes use of parametric properties of stimulus distributions, likelihood ratios, likelihood-ratio-based decision criteria, and other related constructs. Following the lead of many colleagues in the animal learning literature (Davison & Nevin, 1999; Skinner, 1977; White & Wixted, 1999; Wixted & Gaitan, in press), it is important to note that although these constructs help us to understand and characterize decision criterion learning behavior, I am not arguing that people (or nonhuman animals) possess explicit information about the stimulus distributions, likelihood ratios, or decision criteria. Rather, the position is that people gain information (likely implicitly[1]) about the categorization problem as a function of experience with the stimuli and with the reinforcing consequences of the decisions they make (i.e., their reinforcement history). It is acknowledged that there are likely many different mathematical systems (or algorithms) that can capture the behavioral profile of decision criterion learning, and that the present system is only one. Even so, this particular theoretical approach does provide insight into categorization behavior, and more importantly leads to specific testable predictions.

Much of the research in cognitive psychology uses college-age individuals who perform some task for a few hundred trials. Different groups of individuals often participate in different conditions of the experiment, and statistical analyses based on analyses of variance are used. The present approach is different. In line with many colleagues who study animal learning, participants complete several thousand trials and participate in all conditions of the experiment. One advantage of this approach is that there is more control over the participant's reinforcement history, and thus its effects on behavior are more easily identified. By examining a wide range of experimental conditions (i.e., reinforcement histories) within observers, behavioral profiles can be directly compared across reinforcement histories. A second advantage of this approach is that it focuses on the individual. This turns out to be important in category learning because it is often (if not always) the case that behavioral profiles generated by averaging across a large number of participants are not indicative of individual profiles. In fact, Maddox (1999; see also Ashby, Maddox, & Lee, 1994; Estes, 1956; Maddox & Ashby, 1998; Smith & Minda, 1998) showed that averaging alters the structure of categorization data in such a

[1] The distinction between implicit and explicit memory and the importance of each system in category learning have received much attention in the past few years. The current thinking is that there are at least two category learning systems, and that one relies predominantly on explicit memory processes, whereas the other relies predominantly on implicit memory processes (Ashby & Ell, 2001, 2002; Ashby, Maddox, & Bohil, in press). The neurobiological basis of these systems is an area of active research (Filoteo, Maddox, & Davis, 2001a, 2001b; Knowlton, Mangels, & Squire, 1996; Maddox & Filoteo, 2001; see Ashby & Ell, 2001, for a review), and is one in which animal learning theorists are making a major contribution.

way that the correct model of individual performance might provide a poor account of averaged performance, and even worse, an incorrect model of individual performance might provide an excellent account of averaged performance. A third advantage is that a model-based approach to data analyses can be utilized in which we develop a series of nested models that are applied simultaneously to the data from all experimental conditions separately for each observer. Each model makes different assumptions about the effects of the various reinforcement histories on the behavioral profiles observed in each condition.

The second section of this paper outlines the basic properties of many natural categories and offers the normal distribution as a good model of natural categories. The third section provides an overview of the task used to study decision criterion learning. The fourth section formalizes the behavior of the optimal classifier for normally distributed categories, and reviews briefly a signal-detection-based model of categorization, called *decision-bound theory*. This theory, developed by Ashby and Maddox (1993; Maddox & Ashby, 1993), assumes that observers attempt to use the same strategy as the optimal classifier but fail because of inherent suboptimalities in perceptual and cognitive processing. The fifth section introduces a unified theory of decision criterion learning that my colleagues and I have proposed and currently are testing. The sixth section reviews a series of studies that examine the generality of decision criterion learning across experimental contexts that include (a) a wide range of category discriminabilities, (b) cases in which incorrect responses lead to point losses, (c) separate and simultaneous base-rate/payoff manipulations, and (d) feedback based on the objectively correct response or on the response of the optimal classifier. In each case the main empirical results are reviewed and implications for the validity of the unified theory are discussed.

THE NORMAL DISTRIBUTION AS A MODEL OF NATURAL CATEGORIES

To examine rigorously the optimality of categorization performance, one must first identify the basic properties of most everyday categorization problems and build these properties into the experimental categories used in the laboratory. Although no set of properties is common to all natural-environment categories, many such categories have the following six properties. First, the stimulus dimensions typically are continuous valued rather than binary valued. For example, the loudness of the leaves rustling varies continuously. Second, most categories contain a large, often infinite, number of exemplars. For example, there are many levels of loudness of the leaves rustling that are characteristic of a hunter approaching. Third, many categories have a graded structure in which the exemplars are symmetrically and unimodally distributed around some prototype, or at the very least there is evidence that people make this assumption (e.g., Fried & Holyoak, 1984). Fourth, categories generally overlap, meaning that perfect performance is impossible. For example, no matter how advanced a deer's auditory capabilities, there will be times when it continues feeding even though a hunter caused the leaves to rustle and times when it runs when only the wind caused the leaves to rustle. Category overlap is directly related to the predictability or discriminability of the stimulus dimension for correct categorization. For example, the loudness of the leaves rustling might distinguish between a hunter approaching or not approaching when the wind is calm but not during a windstorm. Throughout this article, the term *category discriminability* will be used to refer to the standardized distance between category means, also called *category* d' (Green & Swets, 1967). The more discriminable two categories are along a particular dimension, the larger is d'. Fifth, the category base rates often differ. For example, during hunting season the sound of leaves rustling is more likely to be caused by an approaching hunter than by some harmless source, whereas during the off-season the sound of leaves rustling is less likely to be caused by an approaching hunter. Finally, the benefits associated with correct categorization decisions might differ and the costs associated with incorrect categorization decisions might differ. For example, the benefit of correctly categorizing rustling leaves as indicative of an approaching hunter probably exceeds that of correctly categorizing rustling leaves as not indicative of an approaching hunter, insofar as the former saves the deer's life. Similarly, the cost of falsely categorizing rustling leaves as not indicative of an approaching hunter will exceed that of falsely categorizing rustling leaves as indicative of an approaching hunter, because the former could lead to the deer's death. In general terms, such costs and benefits make up the elements of the *payoff matrix*.

A model that possesses the first four properties is the normal distribution. It is a continuous-valued, unimodal, and symmetric distribution. It contains an infinite number of exemplars that overlap with exemplars from other categories. With normally distributed categories, it is also straightforward to manipulate category discriminability (d'), and to manipulate the category base rates and payoffs associated with each categorization response. The studies reviewed in this article all used normally distributed overlapping categories that were composed of a large number of continuous-valued stimuli within the framework of a perceptual categorization task (also called the *randomization technique*; Ashby & Gott, 1988).

THE PERCEPTUAL CATEGORIZATION TASK

In a typical perceptual categorization task, the experimenter defines two normally distributed categories. On each trial one category is chosen at random in accord with the base rates. A single stimulus is selected at random from the chosen category and is presented to the observer until he or she responds. Once the observer responds, feedback is provided. Following a short intertrial interval, the next stimulus is presented. An example is provided in Figure 1. In this example, the categories are univariate normally distributed, and the stimulus is a bar that varies in height across trials, with ''short'' bars being generally indicative of Category A and ''tall'' bars generally being indicative of Category B. In many of our studies, instructions indicate that the observer is performing a simulated medical diagnosis task in which the height of the bar denotes the outcome of some continuous-valued medical test. In the top portion of Figure 1, a bar height sampled from Category A is presented to the observer. The observer studies the bar and responds ''A'' or ''B.'' The observer is informed of the ''potential gain'' available for the correct response (here, 3 points). Thus, an A response will earn the observer 3 points, with the change as denoted by the ''gain'' in the feedback display. In addition, the ''total gain'' and the ''total potential gain'' earned so far in the experiment are displayed (note that, prior to the current trial, the observer in Figure 1 had 10 points). A B response would earn the observer 0 points. The bottom portion of Figure 1 shows the same information for a stimulus sampled from Category B. Note that in this case the observer receives only 1 point for a correct B response, whereas the same observer received 3 points for a correct A response. Note also that incorrect responses receive 0 points. This is an example of a payoff manipulation in which the gain (or benefit) for a correct A response is three times larger than the benefit for a correct B response. In other studies the benefits are equal for both categories, but the base rates are manipulated in such a way that three times as many stimuli are sampled from Category A than are sampled from Category B. Manipulations of this type form the foundation of the research program described here.

THE OPTIMAL CLASSIFIER AND DECISION-BOUND THEORY

Optimal Classifier

Because the normal distribution is a good model of many everyday categories, it is reflected in the categories of our experimental paradigm, the perceptual categorization task. The next step toward a rigorous examination of the optimality of categorization performance is to formalize the behavior of the optimal classifier with normally distributed categories. The optimal classifier is a hypothetical device that maximizes expected reward (Green & Swets, 1967). Consider the situation facing the deer in the forest that must decide whether a hunter is ''approaching'' or ''not approaching'' based on the loudness of the leaves rustling. Denote ''not approaching'' as Category A, ''approaching'' as Category B, and the loudness of the leaves

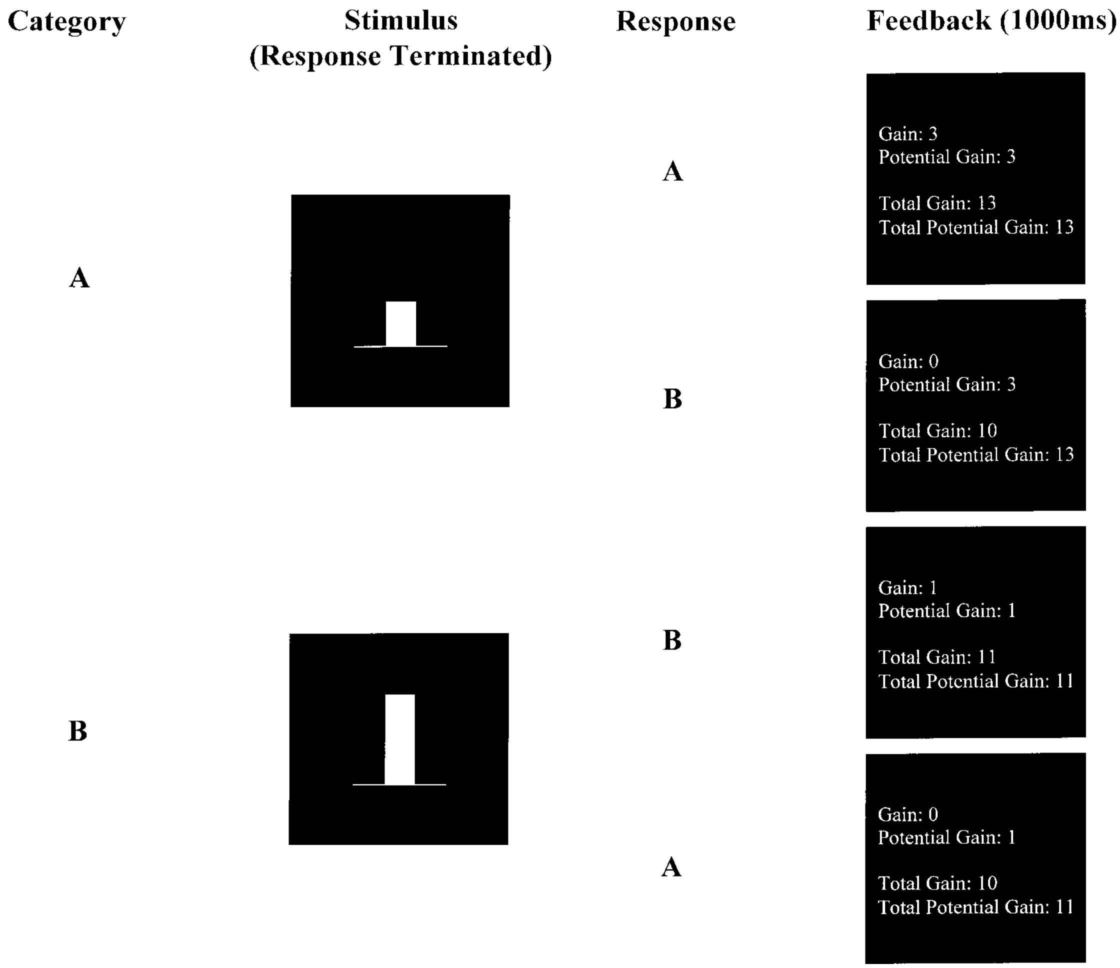

Fig. 1. Timing of a typical trial and hypothetical feedback displays for the perceptual categorization task.

rustling as x. Assume that x is normally distributed for each category with means μ_A and μ_B, and the standard deviations σ_A and σ_B, as shown in Figure 2a. The optimal classifier perfectly records the loudness, denoted as x. In other words, given a fixed physical input, the optimal classifier will show no variability in the perceptual representation. The optimal classifier has perfect knowledge of the form of the category distributions, and the parameters that describe the distribution. This information is used to compute the likelihood ratio of the two category distributions, called the *optimal decision function*,

$$l_o(x) = f(x|\mathrm{B})/f(x|\mathrm{A}), \qquad (1)$$

where $f(x|i)$ is the likelihood of test result x for category i. The likelihood ratio will be greater than one when the likelihood of test result x, given Category B, is larger than the likelihood given Category A. The likelihood ratio for any value of x will be affected by the category discriminability, d'. Three levels of category d' (1.0, 2.2, and 3.2) are depicted in Figure 2.

The optimal classifier has perfect knowledge of the category base rates and the costs and benefits associated with correct and incorrect responses. In other words, the optimal classifier knows the probability that a hunter will be approaching or not approaching. The optimal classifier also knows the benefit of a correct "not approaching" decision, the benefit of a correct "approaching"

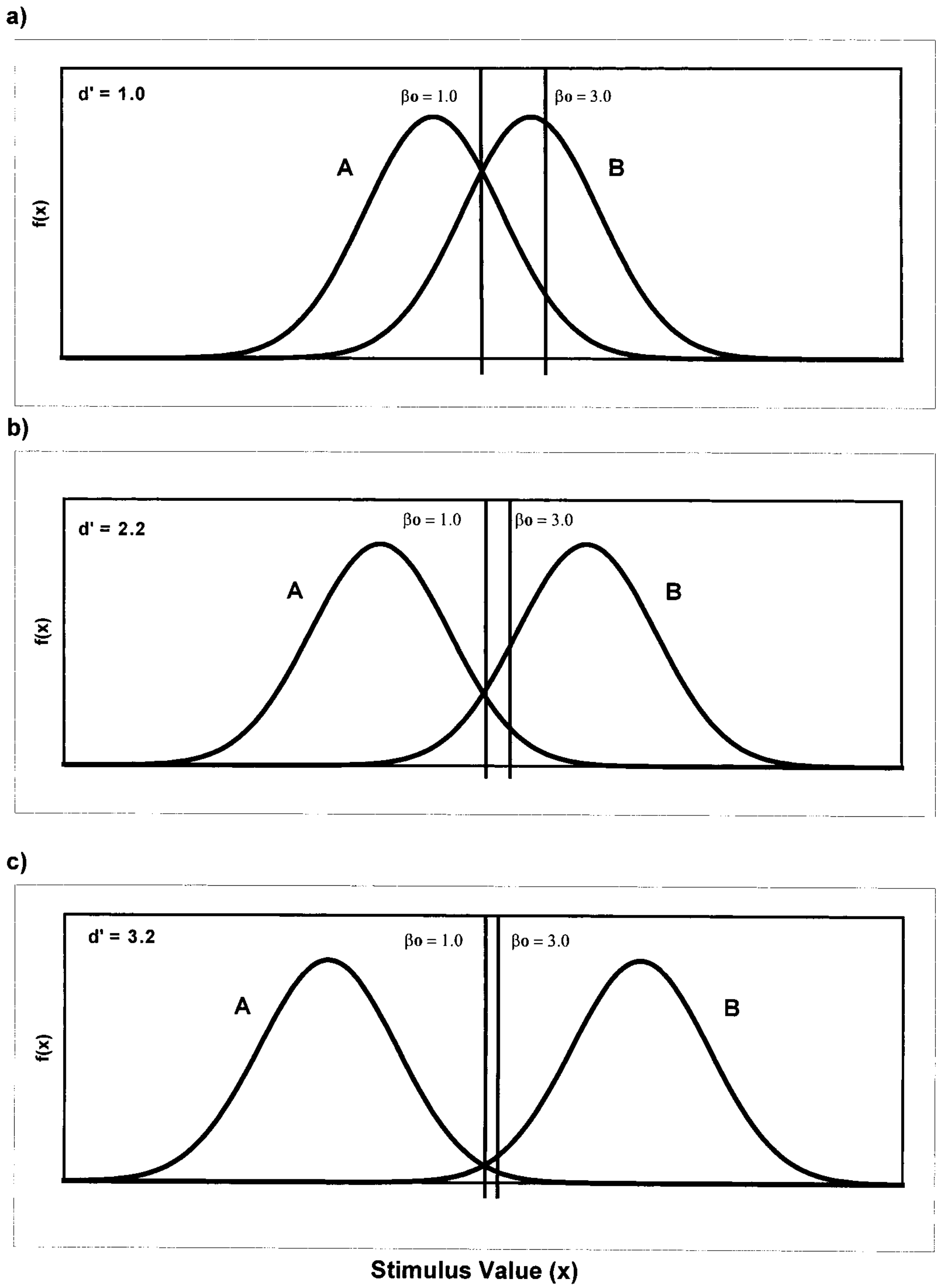

Fig. 2. Hypothetical distributions for Categories A and B when category discriminability (d') is equal to (a) 1.0, (b) 2.2, and (c) 3.2. The $\beta_o = 1$ decision criterion denotes the criterion that is optimal when the base rates are equal and the payoff matrix is symmetric. This is also referred to as the equal likelihood decision criterion. The $\beta_o = 3$ decision criterion denotes the criterion that is optimal when there is a 3:1 base-rate ratio or when the payoff matrix is asymmetric with a 3:1 cost–benefit ratio.

decision, the cost of an incorrect "not approaching" decision, and the cost of an incorrect "approaching" decision. The base rates and the costs and benefits that make up the elements of the payoff matrix are used to determine the value of the *optimal decision criterion*:

$$\beta_o = [P(\mathrm{A})/P(\mathrm{B})] \times [(V_{aA} - V_{bA})/(V_{bB} - V_{aB})], \quad (2)$$

where $P(i)$ is the base-rate (i.e., prior) probability for category i, V_{aA} and V_{bB} are the benefits associated with correct responses, and V_{bA} and V_{aB} are the costs associated with incorrect responses (with lowercase letters denoting response and uppercase letters denoting categories). The optimal classifier uses $l_o(x)$ and β_o to construct the optimal decision rule:

$$\text{If } l_o(x) > \beta_o, \text{ then respond B, otherwise respond A.} \quad (3)$$

Several points are in order. First, note that the Equation 3 decision rule is deterministic, because all stimulus values that fall into one region [where $l_o(x) < \beta_o$] elicit Response A, and all stimulus values that fall into the other region [where $l_o(x) > \beta_o$] elicit Response B. The partition between the regions [where $l_o(x) = \beta_o$] is called the *optimal decision bound*. In the univariate case, the optimal decision bound is simply a point along the stimulus dimension. In the multivariate case, the optimal decision bound is either linear or quadratic (see Ashby & Maddox, 1990, 1992, for examples). Second, note that when the base rates are equal and the cost–benefit differences are equal [i.e., $P(\mathrm{A}) = P(\mathrm{B})$ and $(V_{aA} - V_{bA}) = (V_{bB} - V_{aB})$], $\beta_o = 1.0$, and the response associated with the most likely category is given. Third, note also that situations exist for which a base-rate and payoff manipulation can have identical effects on the optimal decision criterion. For example, suppose the category "hunter not approaching" is three times more likely than the category "hunter approaching," and the payoff matrix is unbiased [i.e., if $P(\mathrm{A}) = 3P(\mathrm{B})$ and $(V_{aA} - V_{bA}) = (V_{bB} - V_{aB})$], or suppose the cost–benefit difference for "hunter not approaching" is three times larger than for "hunter approaching" and the base rates are equal [i.e., when $(V_{aA} - V_{bA}) = 3(V_{bB} - V_{aB})$ and $P(\mathrm{A}) = P(\mathrm{B})$], then $\beta_o = 3.0$ (see Figure 2). In this case, the optimal classifier will make a "hunter not approaching" decision unless the likelihood of the "hunter approaching" is at least three times larger than the likelihood of the "hunter not approaching." Finally, when base rates and payoffs are manipulated simultaneously, the optimal decision criterion can be derived from an independent combination of the separate base-rate and payoff decision criteria. This is seen more clearly in a mathematically equivalent formulation of Equation 2 in which the natural log is applied to both sides, yielding

$$\log \beta_o = \log[P(\mathrm{A})/P(\mathrm{B})] + \log[V_{aA} - V_{bA})/(V_{bB} - V_{aB})]. \quad (4)$$

Notice that $\log\beta_o$ is determined completely by the sum of an independent base-rate and payoff term. This is referred to as the *independence assumption* of the optimal classifier.

This example illustrates that optimal categorization performance that maximizes expected reward (e.g., length of life) requires knowledge of the distributional properties of the categories, such as category discriminability, the category base rates, and the costs and benefits associated with different responses. For example, whether the deer concludes that a hunter "is approaching" or "is not approaching" based on the loudness of the leaves rustling depends on how reliably the loudness of the leaves rustling indicates a hunter approaching (category d' information), the likelihood of a hunter during that time of the year (category base-rate information), how hungry the deer is, whether a fawn is present, and so forth (cost–benefit information). To reiterate a point made earlier, I am not arguing that organisms have explicit knowledge of these details. Rather, my position is that experience with these stimuli and with the consequences of the organism's decisions leads the organism to behave in a manner that can often be captured by the optimal classifier or by models that incorporate reasonable suboptimalities. The approach taken in the research reviewed below was to expose observers to a wide range of category distribution, base-rate, and payoff conditions, and to compare observers' performance with that of the optimal classifier. When performance did not match that of the optimal classifier, the aim was to develop psychologically meaningful hypotheses regarding the locus of these suboptimalities, and to test them within the framework of a statistically rigorous model-based approach. The foundation of this model-based approach is provided by a signal-detection-based model of categorization, called *decision-bound theory* (Ashby, 1992a; Ashby & Maddox, 1993; Maddox & Ashby, 1993), which is reviewed next.

Decision-Bound Theory

The optimal classifier decision rule (Equation 3) has been rejected as a model of human performance, but in many cases, performance approaches that of the optimal classifier as the observer gains experience with the task. Ashby and colleagues argued that the observer attempts to respond in accord with the optimal classifier, but fails because of various suboptimalities in perceptual and cognitive processing (Ashby, 1992a; Ashby & Lee, 1991; Ashby & Maddox, 1993, 1994; Ashby & Townsend, 1986; Maddox & Ashby, 1993). They proposed a series of decision bound models to test specific hypotheses about the locus of performance suboptimalities. Two suboptimalities inherent in humans and other organisms are *perceptual* and *criterial noise*. Perceptual noise exists because there is trial-by-trial variability in the perceptual information associated with each stimulus. Assuming a single perceptual dimension is relevant, the observer's perception of stimulus *i*, on any trial, is given by $x_{pi} = x_i + e_p$, where x_i is the observer's mean perception, and e_p is a random variable that represents the effects of perceptual noise. At the cognitive level, there is trial-by-trial variability in the observer's memory for the decision criterion (termed criterial noise). Because of criterial noise, the decision criterion used on any trial is given by $\beta_c = \beta + e_c$, where β is the observer's average decision criterion, and e_c is a random variable that represents the effects of criterial noise (assumed to be univariate normally distributed).

Because perceptual and criterial noise exist, the observer cannot attain the level of performance reached by the optimal classifier (i.e., cannot maximize expected reward). Even so, decision-bound theory assumes that the observer attempts to use the same strategy as the optimal classifier, but with less success due to the effects of perceptual and criterial noise (and other possible suboptimalities). Hence, the simplest decision-bound model is the *optimal decision-bound model*. The optimal decision-bound model is identical to the optimal classifier (Equation 3) except that perceptual and criterial noise are incorporated into the decision rule. Specifically,

$$\text{If } l_o(x_{pi}) > \beta_o + e_c,$$

$$\text{then respond B, otherwise respond A.} \tag{5}$$

It is important to note that the optimal decision-bound model often predicts performance that is very nearly optimal. For example, in many cases, perceptual noise will be small, and the perceptual representation will be close to veridical. In addition, experience with a task and certain types of decision criteria decrease the magnitude of criterial noise. Despite this fact, it is important to acknowledge these inherent sources of noise and to account for them within theories of categorization.

Before continuing, a few words are in order regarding the relations between the decision rules of the optimal classifier (Equation 3) and the optimal decision-bound model (Equation 5). As outlined above, the optimal classifier decision rule is deterministic. Specifically, the same response is always given to the same stimulus. In the learning literature this is commonly referred to as a maximizing strategy (Herrnstein, 1961, 1970; Herrnstein & Heyman, 1979; Williams, 1988). The optimal decision-bound model decision rule, on the other hand, is not deterministic, because the same response is not always given to the same stimulus, which implies some level of matching. It is deterministic, however, in the sense that the probability of responding B is equal to 1 when $l_o(x_{pi}) > \beta_o + e_c$. Thus, the optimal decision-bound model can predict responding that ranges from maximizing (when perceptual and criterial noise are 0) through various levels of matching (e.g., over- or undermatching), depending on the magnitude of the perceptual and criterial noise. Issues of matching and maximizing in category learning are addressed by Ashby and Maddox (1993).

In the mid-1990s I began a research program to examine decision criterion learning

(Equation 2). This research is unique because it bridges the gap between traditional categorization studies that focus on processes involved in category structure learning and decision-making studies that focus on processes involved in base-rate and payoff learning by allowing both issues to be examined within a single unified theoretical framework.

A UNIFIED THEORY OF DECISION CRITERION LEARNING IN PERCEPTUAL CATEGORIZATION

This section describes the observers and the basic design of the experiments, then briefly reviews two important empirical results that led to the development of the unified theory. Finally, the theory is described, and four studies conducted to test the unified theory are reviewed.

Observer and Task Specifics

All of the studies reviewed below used the perceptual categorization task described in Figure 1. Each observer participated in all conditions of each experiment (i.e., a within-observer design was used). Depending on the study, this required the observer to complete 4 to 15 experimental sessions and several thousand trials. In light of this fact, all observers were recruited from the university community through advertisements and were paid for their participation. Observers were instructed to maximize the number of points earned in each experimental condition, and were told that their point totals would be converted into money that would constitute their compensation for participating. A typical experimental session lasted approximately 1 hr, and the observer generally completed 400 to 800 trials in a session. All of the studies used univariate normally distributed categories and the height of a bar as the stimulus.

Because the focus was on decision criterion learning that is primarily affected by base-rate and payoff manipulations (Equation 2), and was not on decision function learning that is primarily affected by the category distributions (Equation 1), it was important to ensure that observers had accurate knowledge of the category structures before being exposed to base-rate and payoff manipulations. To achieve this goal, all of the studies reviewed below contained at least one session (400 to 800 trials) at the beginning of the study and 60 trials at the beginning of each experimental condition in which no base-rate or payoff manipulation was present. This is referred to as the baseline condition. Observers were trained in the baseline condition until performance reached a rigid accuracy-based performance criterion, and the optimal decision bound model (Equation 5) provided the most parsimonious account of the data. This approach ensures accurate knowledge of the category structures, allowing us to assume that the observer is using the optimal decision function [i.e., $l(x) = l_o(x)$], and minimizes any within-observer carryover effects from one experimental condition to the next.

A Unified Theory of Decision Criterion Learning and a Hybrid Model Framework

In the 1960s and 1970s several experiments examined decision criterion learning under unequal base-rate or payoff conditions (e.g., Green & Swets, 1967; Healy & Kubovy, 1981; Kubovy & Healy, 1977; Lee & Janke, 1964, 1965; Lee & Zentall, 1966; Ulehla, 1966). Two robust findings emerged from this work. First, comparisons of the optimal decision criterion with the observer's decision criterion suggested that observers used a criterion that was more conservative than the optimal decision criterion. For example, if the base rates or payoffs were such that $\beta_o = 3$, then observers tended to use a β between 1 and 3. This was termed *conservative cutoff placement*, because the decision criterion was not shifted far enough toward the optimal value. Second, observers' decision criterion estimates were closer to the optimal value when base rates as opposed to payoffs were manipulated, even when the optimal decision criterion was identical across base-rate and payoff conditions. Several explanations for these results have been offered in the literature and are reviewed elsewhere (Healy & Kubovy, 1981; Maddox & Bohil, 2000), but none has been generally accepted.

The first few studies conducted in our laboratory were designed to replicate and extend the previous research (Bohil & Maddox, 2001; Maddox, 1995; Maddox & Bohil, 1998a, 1998b, 2000) using decision-bound theory. Although useful, one weakness of the decision-bound theoretical approach to decision criterion learning used in these studies is that no mechanisms were postulated or formalized to guide decision criterion placement. Rather, the decision criterion, β, was freely estimated from the data.

In this section I outline a formal theory of decision criterion learning, and a model based instantiation. The idea is to use decision bound theory as the basic modeling framework, but to supplement the model by postulating psychologically meaningful mechanisms that guide decision criterion placement. In other words, instead of simply estimating the decision criterion value directly from the data, I propose two mechanisms that constrain the decision criterion value. These mechanisms are consistent with the laboratory conditions under which conservative cutoff placement and better decision criterion learning in base-rate over payoff conditions have been observed.

Flat-maxima hypothesis. The first mechanism is based on the *flat-maxima hypothesis* (Busemeyer & Myung, 1992; von Winterfeldt & Edwards, 1982) and was developed to account for the finding that observers tend to use a criterion that is more conservative than the optimal decision criterion when base rates or payoffs are manipulated (conservative cutoff placement). As suggested by many researchers, suppose that the observer adjusts the decision criterion based (at least in part) on the change in the rate of reward, with larger changes in rate being associated with faster, more nearly optimal decision criterion learning (e.g., Busemeyer & Myung, 1992; Dusoir, 1980; Erev, 1998; Erev, Gopher, Itkin, & Greenshpan, 1995; Kubovy & Healy, 1977; Roth & Erev, 1995; Thomas, 1975; Thomas & Legge, 1970). To formalize this hypothesis one can construct the *objective reward function.* The objective reward function plots objective expected reward on the *y* axis and the decision criterion value on the *x* axis (e.g., Busemeyer & Myung, 1992; von Winterfeldt & Edwards, 1982). To generate an objective reward function, one chooses a value for the decision criterion and computes the expected reward for that criterion value. This process is then repeated numerous times for different decision criterion values. The expected reward is then plotted as a function of the decision criterion value. Figure 3a displays the objective reward function for category d'

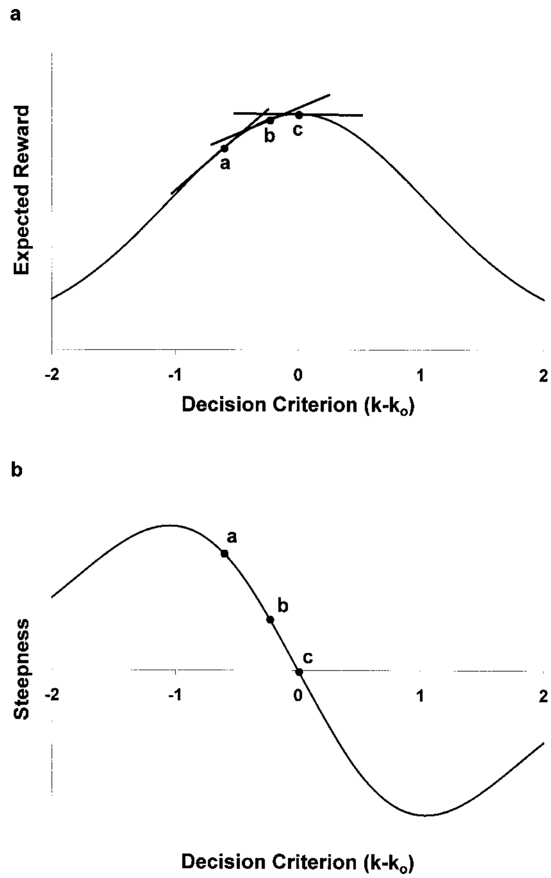

Fig. 3. Panel a: expected reward as a function of the decision criterion (relative to the optimal decision criterion; i.e., $k - k_o$), called the objective reward function for category discriminability, $d' = 1.0$. The three lines are the tangent lines at Points a, b, and c on the objective reward function that denote the derivative or steepness of the objective reward function at each point. Panel b: steepness of the objective reward functions from Panel a along with the three points highlighted in Panel a.

$= 1.0$. Specifically, Figure 3a plots expected reward as a function of the deviation between the decision criterion (β) and the optimal decision criterion (β_o) standardized by category d'. This $k - k_o$ measure $= \log(\beta)/d' - \log(\beta_o)/d' = \log(\beta/\beta_o)/d'$ is the ratio of the observed and optimal decision criterion standardized by category d'. Notice that for large deviations from the optimal decision criterion, the expected reward is small, and as the deviation from the optimal decision criterion decreases, the expected reward increases. Notice also that when the deviation from optimal is zero (i.e., when the decision criterion is the optimal decision criterion), expected reward is maximized.

The derivative of the objective reward function at a specific $k - k_o$ value determines the change in the rate of expected reward for that $k - k_o$ value; the larger the change in the rate, the "steeper" the objective reward function at that point. Derivatives for three $k - k_o$ values are denoted by the three tangent lines denoted as a, b, and c in Figure 3a. Notice that the slope of each tangent, which defines the derivative, decreases as the deviation from the optimal decision criterion decreases (i.e., as we go from Point a to b to c). In other words, the change in the rate of reward or steepness declines as the decision criterion approaches the optimal decision criterion. Figure 3b plots the relationship between the steepness of the objective reward function (i.e., the derivative at several $k - k_o$ values) and $k - k_o$. The three derivatives denoted in Figure 3a are highlighted in Figure 3b. If the observer adjusts the decision criterion based on the change in the rate of reward (or steepness of the objective reward function), as described above, then steeper objective reward functions should be associated with more nearly optimal decision criterion values, because only a small range of decision criterion values around the optimal value have nearly zero derivatives (or small steepness values). Flat objective reward functions, on the other hand, will lead to less optimal decision criterion placement because a larger range of decision criterion values around the optimal value have derivatives near zero. Interestingly, nearly all the work conducted in the 1960s and 1970s used tasks in which the category discriminability (d') was 1.0. The objective reward function is shallow for $d' = 1.0$, which might explain the prevalence of conservative cutoff placement (see Figure 3a). As we will see shortly, the flat-maxima hypothesis makes strong predictions about the effects of several environmental factors such as category d', the base-rate/payoff ratio, and various linear transformations of the payoff matrix entries on decision criterion placement.

It is important to note that the flat-maxima hypothesis applies only to learning of the *reward-maximizing* decision criterion. As outlined shortly, the observed decision criterion is assumed to be a weighted average of the reward- and accuracy-maximizing decision criteria. Although the flat-maxima hypothesis offers an explanation for the prevalence of conservative cutoff placement, it does not offer an explanation for the robust finding that the observed decision criterion is closer to optimal in base-rate, compared to unequal payoff, conditions.

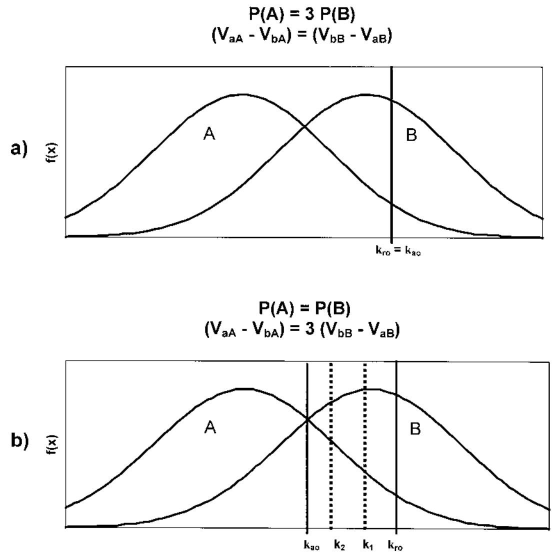

Fig. 4. Schematic illustration of the COBRA hypothesis. See text for details.

Competition-between-reward-and-accuracy-maximization (COBRA) hypothesis. The second mechanism assumed to influence decision criterion placement is based on Maddox and Bohil's (1998a) COBRA hypothesis and was developed to account for the finding that observers show more nearly optimal decision criterion placement in unequal base-rate conditions than in unequal payoff conditions. COBRA postulates that observers, although they attempt to maximize expected reward (consistent with instructions, and monetary compensation contingencies), also place importance on accuracy maximization. In other words, in the observer's reinforcement history both goals have been associated with primary reinforcement, so even though they are instructed to maximize reward in the laboratory, they cannot because of conditioned reinforcing outcomes. Consider the univariate categorization problems depicted in Figure 4. Panel a displays a 3:1 base-rate condition, and Panel b displays a 3:1 payoff condition. As suggested by Equation 2, expected reward is maximized in both cases by using the optimal reward-maximizing decision criterion, $k_{ro} = \log(\beta_{ro})/d' = \log(3)/d'$. Thus, an observer who attempts to maximize expected reward should use the same decision criterion in both conditions. However, whereas the accuracy- and reward-maximizing decision criteria are the same in the 3:1 base-rate condition (i.e., $k_{ao} = k_{ro}$), they are different in the 3:1 payoff condition [$k_{ao} = \log(\beta_{ao})/d' = \log(1)/d'$]. (When base rates are equal, it is always the case that the accuracy-maximizing decision criterion $\beta_{ao} = 1$.) When base rates are manipulated, accuracy and reward can be maximized simultaneously, because $k_{ro} = k_{ao}$, but when payoffs are manipulated, both goals cannot be achieved simultaneously because $k_{ro} \neq k_{ao}$. An observer who places importance (or weight) on both goals will use a decision criterion intermediate between the accuracy- and reward-maximizing decision criteria in the payoff condition, and thus will show more conservative cutoff placement in the payoff condition than in the base-rate condition. To instantiate this hypothesis we assume a simple weighting function, $k = wk_a + (1 - w)\ k_r$, where w ($0 \leq w \leq 1$) denotes the weight placed on expected accuracy maximization. Other weighting schemes are possible. For example, instead of generating an intermediate decision criterion, it is possible that the two decision criteria compete on each trial for the opportunity to generate the categorization response (for related proposals, see Ashby, Alfonso-Reese, Turken, & Waldron, 1998). The current approach is simple to instantiate and has met with reasonable success (Maddox & Dodd, 2001). This weighting function results in a single decision criterion that is intermediate between that for accuracy maximization and that for reward maximization. For example, in Figure 4b, k_1 denotes a case in which $w < .5$, whereas k_2 denotes a case in which $w > .5$.

Framework for a hybrid model. Maddox and Dodd (2001) developed a hybrid model of decision criterion learning that incorporated both the flat-maxima and COBRA hypotheses. Specifically, the model assumes that the decision criterion used by the observer to maximize expected reward (k_r) is determined by the steepness of the objective reward function (see Figure 3). A single steepness parameter is estimated from the data that determines a distinct decision criterion in every condition for which the steepness of the objective reward function differs. (As will be shown shortly, several different experimental manipulations, e.g., category discriminability, have strong effects on the steepness of the objective reward function and thus offer powerful tests of the flat-maxima hypothesis.) The COBRA hypothesis is instantiated in the hybrid model by estimating the accuracy weight, w, from the data. To facilitate the development of each model, consider the following equation that determines the decision criterion used by the observer on condition i trials (k_i):

$$k_i = wk_a + (1 - w)k_r. \tag{6}$$

When base rates are manipulated, the observer's estimate of the reward-maximizing decision criterion, derived from the flat-maxima hypothesis, is also the best estimate of the accuracy-maximizing decision criterion, resulting not in competition but simply in use of the reward-maximizing decision criterion. When payoffs are manipulated, on the other hand, the reward- and accuracy-maximizing decision criteria are different. Fortunately, by pretraining each observer on the category structures in the baseline condition (described earlier), we are essentially pretraining the accuracy-maximizing decision criterion. This criterion is then entered into the weighting function along with the observer's estimate of the reward-maximizing decision criterion to determine the criterion used on each trial.

Model details and general nested model fitting procedure. Because all of the studies conducted by my colleagues and me used within-observer designs and examined decision criterion learning across several blocks of trials, we could apply the hybrid model framework simultaneously to the data from all experimental conditions separately for each observer and for each block of trials. The model parameters were estimated using maximum likelihood procedures (for details, see Ashby, 1992b; Maddox & Dodd, 2001; Wickens, 1982). In every application of the hybrid model framework to individual observer decision criterion learning data, we began by applying four "base" models, each of which makes different assumptions about the k_r and

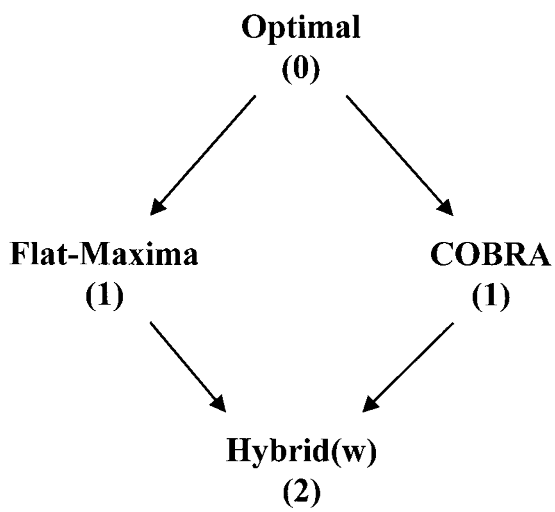

Fig. 5. Nested relation among the four base versions of the hybrid model. These models were applied to all studies reviewed in this article. The number in parentheses denotes the number of free parameters. The arrows point to a more general model. See text for details.

w values. These four base models have a nested structure, which means that a more specific or "nested" model can be obtained from a more general model by setting some of the parameters of the more general model to constants. One advantage of a nested modeling approach is that rigorous statistical tests can be used to determine whether the additional parameters of a more general model provide a statistically significant improvement in fit over a more specific, nested model. The details of this nested model testing procedure are outlined elsewhere (e.g., Ashby, 1992b; Maddox & Dodd, 2001; Wickens, 1982) and will not be elaborated here.

The nested structure of the four base models is presented in Figure 5. The number of free parameters is presented in parentheses. The arrows point to the more general model. Models at the same level have the same number of free parameters. The *optimal model* assumes that the reward-maximizing decision criterion is optimal (i.e., $k_r = k_o$), and that there is no competition between reward and accuracy maximization (i.e., $w = 0$). This model instantiates neither the flat-maxima nor the COBRA hypothesis. The *flat-maxima model* assumes that the reward-maximizing decision criterion (k_r) is determined by the steepness of the objective reward function, and that there is no competition between accuracy and reward maximization (i.e., $w = 0$). A single steepness parameter is estimated from the data. This single steepness parameter determines a different decision in every condition for which the steepness of the objective reward function differs. This model contains the optimal model as a special case in which the steepness value is equal to zero (i.e., $k_r = k_o$), and instantiates the flat-maxima hypothesis but not the COBRA hypothesis. The *COBRA model* assumes that the reward-maximizing decision criterion is optimal (i.e., $k_r = k_o$), but permits competition between reward and accuracy maximization by estimating the Equation 6 *w* parameter from the data. This model contains the optimal model as a special case in which $w = 0$, and instantiates the COBRA hypothesis but not the flat-maxima hypothesis. The *hybrid (w) model* instantiates both the flat-maxima and the COBRA hypotheses. It assumes that k_r is determined by the steepness of the objective reward function, and that there is a competition between accuracy and reward maximization. This model includes the previous three models as special cases.

Across a large range of experimental contexts, the most common finding was for the hybrid model to provide a statistically significant improvement in fit over the two models that instantiate only one component of the unified theory (i.e., the flat-maxima and COBRA models), and over the optimal model that instantiates neither. In addition, the fit of the hybrid model was generally good, usually accounting for over 90% of the responses from each individual observer. Depending on the experimental factors of interest, more general models were also applied to the data. In all of these cases, the more general models were constructed by including different accuracy weight parameters for different experimental conditions. In the review that follows, only the most informative models will be discussed. I turn now to a review of recent work that examined decision criterion learning under a range of experimental conditions.

EXTENDING THE RANGE OF EXPERIMENTAL DECISION CRITERION LEARNING CONTEXTS

Category Discriminability Effects on Decision Criterion Learning

Most previous studies of decision criterion learning focused on cases in which category discriminability $d' = 1.0$. Observers showed conservative cutoff placement in base-rate and payoff conditions, with the magnitude of conservative cutoff placement being larger in payoff than in base-rate conditions. Because the objective reward function is relatively flat for $d' = 1.0$, one possibility is that some of the performance suboptimality observed in these studies was due to poor learning of the reward-maximizing decision criterion. It is possible that better decision criterion learning might result if the objective reward function was steeper. Thus, it is of interest from an empirical standpoint, and from the standpoint of our unified theory, to extend the range of category discriminabilities investigated.

Figure 6a displays the objective reward functions for three levels of category d': 1.0, 2.2, and 3.2. Figure 6b plots the relation between the steepness for each objective reward function (i.e., the derivatives for each objective reward function) and $k - k_o$. The tangent lines in Figure 6a labeled 1, 2, and 3 denote the $k - k_o$ value associated with the same fixed steepness value for $d' = 1.0$, 2.2, and 3.2, respectively. The horizontal line on Figure 6b denotes the same fixed nonzero steepness value, and the vertical lines denote the associated $k - k_o$ values for each category d'. Notice that for this fixed nonzero steepness the deviation between the decision criterion and the optimal value, $k - k_o$, differs systematically across category d' conditions in such a way that the decision criterion, k, is closest to the optimal value, k_o, for category $d' = 2.2$, is farthest from optimal for $d' = 1.0$, and is intermediate for $d' = 3.2$. If the observer adjusts the reward-maximizing decision criterion based on the change in the rate of reward (or steepness of the objective reward function), as suggested by the flat-maxima hypothesis, then performance should be closest to optimal when category $d' = 2.2$, farthest from optimal when category $d' = 1.0$, and intermediate when category $d' = 3.2$.

Maddox and Dodd (2001) examined decision criterion learning in a 3:1 base-rate condition, two 3:1 payoff no-cost conditions, and two 3:1 payoff cost conditions at each of three category discriminabilities ($d' = 1.0$, 2.2, and 3.2) for a total of 15 conditions. In the 3:1 payoff no-cost conditions, the cost of an incorrect response was set to zero. In the 3:1 payoff cost conditions, the observer lost points for an incorrect response. Data from the two no-cost conditions and from the two cost conditions were collapsed. Each of 6 observers completed six 60-trial blocks in each condition. Figure 7 displays data from a representative observer, showing the deviation from optimal points, defined as (observed points − optimal points)/(optimal points − points for 0% correct). Two key results stand out. First, in line with the flat-maxima hypothesis, performance was closest to optimal

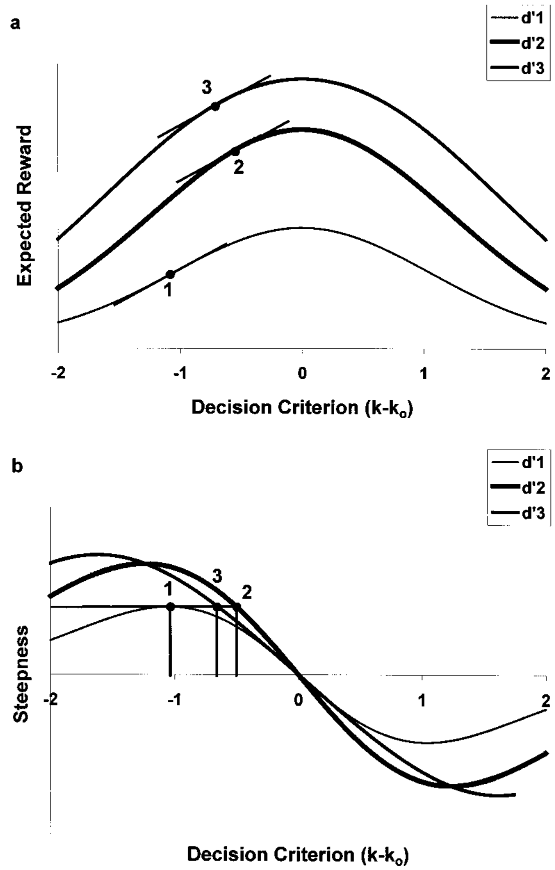

Fig. 6. Panel a: expected reward as a function of the decision criterion (relative to the optimal decision criterion; i.e., $k - k_o$), called the objective reward function for category discriminability, $d' = 1.0$, 2.2, and 3.2. The three lines are tangent lines, one for each d' level, with the same slope and thus the same derivative or steepness. Panel b: steepness of the objective reward functions from Panel a along with the three points highlighted in Panel a. The three points all have identical steepness as denoted by the horizontal line. The resulting decision criterion is closer to the optimal value for $d' = 2.2$, is farthest from optimal for $d' = 1.0$, and is intermediate for $d' = 3.2$.

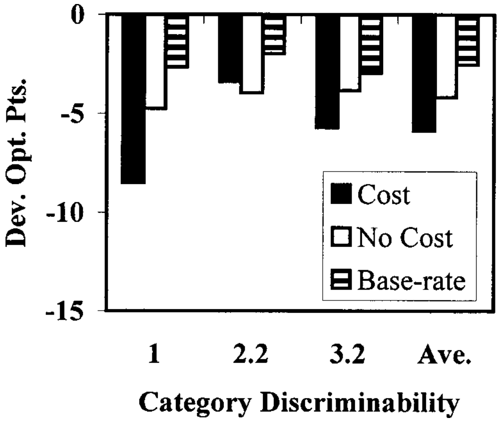

Fig. 7. Deviation from optimal points for a representative observer from Maddox and Dodd (2001).

for $d' = 2.2$, was farthest from optimal for $d' = 1.0$, and was intermediate for $d' = 3.2$. Second, performance was closer to optimal in the base-rate condition than in either payoff condition, as predicted by COBRA. In addition, performance was closer to optimal in the payoff no-cost conditions (when no points were lost for an incorrect response) than in the payoff cost conditions (when points were lost for an incorrect response). The latter finding is important because, drawing upon decision theory, Maddox and Bohil (2000) had speculated that more emphasis might be placed on accuracy maximization (and thus point totals would be farther from optimal) when incorrect responses led to an actual loss of points.

The four base models outlined in Figure 5 were applied simultaneously to the data from all 15 conditions separately by observer and 60-trial block. One additional model, the *hybrid* (w_{cost}; $w_{no\ cost}$) *model,* was developed to test the hypothesis proposed by Maddox and Bohil (2000) that the weight placed on accuracy maximization was greater when incorrect responses led to an actual loss of points. In this model the w_{cost} parameter was applied to the data from the payoff cost conditions, and the $w_{\text{no cost}}$ parameter was applied to the data from the payoff no-cost conditions. The hybrid (w_{cost}; $w_{\text{no cost}}$) model accounted for 91% to 93% of the responses in the data from each observer (92.2% of the responses in the data from the representative observer displayed in Figure 7). It is important to emphasize that the hybrid (w_{cost}; $w_{\text{no cost}}$) model assumes that the observer's reward-maximizing decision criterion was determined from the steepness of the objective reward function. Because of the strong influence of category discriminability on the steepness of the objective reward function (see Figure 6), this model is constrained to predict that the reward-maximizing decision criterion will be closest to optimal for $d' = 2.2$, farthest from optimal for $d' = 1.0$, and intermediate for $d' = 3.2$. The excellent fit of this model provides strong support for the flat-maxima hypothesis as it applies to category discriminability manipulations.

The models were applied separately to each block of trials but the most interesting performance trends were observed early and late in learning. Figure 8a displays the steepness, w_{cost}, and $w_{\text{no cost}}$ values early and late in learning for a representative observer, and Figure 8b displays the same parameters averaged across fits to each individual observer's data. The results can be summarized as follows. First, there was a large decline in the steepness parameter from early to late in the session, suggesting that the observer's estimate of the reward-maximizing decision criterion, k_r, approached the optimal value. Second, the weight placed on accuracy in both the cost and no-cost conditions declined from early to late in the session, suggesting that observers became more willing to sacrifice accuracy in the interest of reward maximization as they gained experience with the task. Finally, both early and late in the session, the weight placed on accuracy was larger in the cost than in the no-cost conditions, suggesting that observers were less willing to sacrifice accuracy when they lost points for incorrect responses.

Payoff Matrix Multiplication and Addition Effects on Decision Criterion Learning

Most previous studies of the effects of unequal payoffs on decision criterion learning examined cases in which the cost of an incorrect response led to no loss of points. This is referred to as a payoff no-cost condition. In the last section, we reviewed a study by Maddox and Dodd (2001) that compared decision criterion learning in payoff no-cost conditions with decision criterion learning in payoff cost conditions. They found poorer decision criterion learning in payoff cost con-

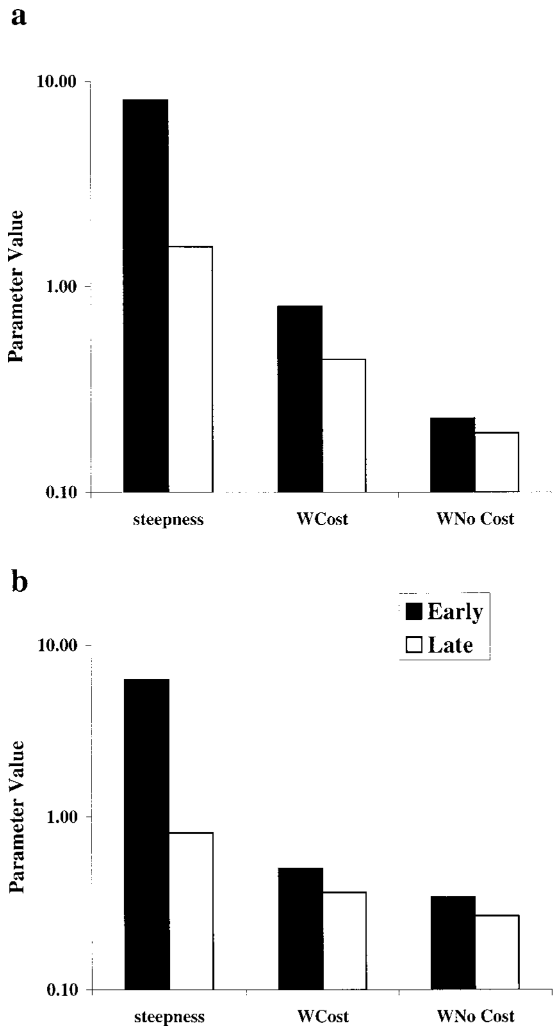

Fig. 8. Panel a: parameter values from the hybrid (w_{cost}; $w_{no\ cost}$) model for a representative observer from Maddox and Dodd (2001). Panel b: parameter values from the hybrid (w_{cost}; $w_{no\ cost}$) model averaged across fits of the model separately to each individual observer. The parameter values are plotted on a log scale for the data early in learning (Blocks 1 and 2) and late in learning (Blocks 5 and 6).

ditions, and suggested that observers might place more weight on accuracy maximization when an incorrect response led to a reduction in the number of points. The accuracy weight parameters estimated from the hybrid model supported this claim. In light of this finding, and because the costs and benefits associated with different natural-environment categorization problems vary widely, an examination of decision criterion learning under a wide range of cost–benefit conditions is in order. Maddox, Dodd, and Bohil (2001)[2] examined decision criterion learning in each of eight 3:1 payoff conditions. Consider the condition denoted shallow/no cost (B) in Table 1. In this condition, $V_{aA} = 4$, $V_{bB} = 2$, $V_{aB} = 1$, and $V_{bA} = 1$. Notice that the shallow/no-cost (A), shallow/cost-LRG, and shallow/cost-LRL conditions can be derived from the shallow/no-cost (B) condition by subtracting 1, 2 and 3, respectively, from all payoff matrix entries (LRG stands for long-run gain and LRL stands for long-run loss, both of which will be elaborated shortly). Each of these conditions is related to the other via a simple additive transformation referred to as *payoff matrix addition (PMA)*. Next take each of the four "shallow" payoff matrices and multiply all entries by six. The resulting four "steep" payoff matrices are displayed in Table 1. Each of the four steep payoff matrices is related to its associated shallow payoff matrix by a simple multiplicative transformation referred to as *payoff matrix multiplication (PMM)*. In the no-cost and cost-LRG conditions, the optimal classifier will start the experiment with 0 points, and will gain points over the course of the experiment. In the cost-LRL conditions, on the other hand, the optimal classifier will start the experiment with 0 points, and will lose points over the course the experiment.

Based on previous research (Maddox & Bohil, 2000; Maddox & Dodd, 2001) the most reasonable prediction is that more weight will be placed on accuracy in the cost conditions than in the no-cost conditions, but no a priori predictions have been offered with respect to the cost-LRG versus cost-LRL comparison. One possibility is that the weight placed on accuracy will be the same in the two cost conditions. A second possibility is that the weight placed on accuracy will be greater when there is a long-run loss of points accrued over the course of the experiment, perhaps because the observer realizes that there is no way to gain points, thus leading to an exclusive focus on accuracy. A third possibility is that less weight will be placed on accuracy in the long-

[2] Maddox, W. T., Dodd, J. L., & Bohil, C. J. (2001). *Payoff matrix multiplication and addition effects on decision criterion learning in simulated medical diagnosis.* Unpublished data.

Table 1

Category cost–benefit conditions from Maddox, Dodd and Bohil (2001). P(A) = Category A base rate. P(B) = Category B base rate. V_{iJ} = the value (cost or benefit) associated with an i response to a stimulus from Category J. LRL = long-run loss. LRG = long-run gain.

	Base rates		Cost benefits				
	P(A)	P(B)	V_{aA}	V_{bA}	V_{bB}	V_{aB}	β_o
Shallow/cost LRL	.50	.50	1	−2	−1	−2	3
Shallow/cost LRG	.50	.50	2	−1	0	−1	3
Shallow/no cost (A)	.50	.50	3	0	1	0	3
Shallow/no cost (B)	.50	.50	4	1	2	1	3
Steep/cost LRL	.50	.50	6	−12	−6	−12	3
Steep/cost LRG	.50	.50	12	−6	0	−6	3
Steep/no cost (A)	.50	.50	18	0	6	0	3
Steep/no cost (B)	.50	.50	24	6	12	6	3

run loss condition because the observer realizes (perhaps implicitly) that no strategy will yield a gain over the course of the experiment, thus leading to less focus on accuracy.

The shallow versus steep distinction is made because the steepness of the objective reward function is affected by PMM. Figure 9a displays the objective reward function for the shallow/no-cost (A) and steep/no-cost (A) payoff conditions used in this study. The objective reward functions for the other three shallow conditions have the same shape as the shallow/no-cost (A) condition, but have different asymptotes. Similarly, the objective reward functions for the other three steep conditions had the same shape as the steep/no-cost (A) condition, but have different asymptotes. Figure 9b plots the relation between the steepness for each objective reward function (i.e., the derivatives for each objective reward function) and $k - k_o$. The horizontal line on Figure 9b denotes a fixed nonzero steepness value, and the vertical lines denote the associated $k - k_o$ values. First, notice that for a fixed nonzero steepness the reward-maximizing decision criterion, k, differed systematically across steep and shallow payoff matrix conditions in such a way that the reward-maximizing decision criterion is closer to the optimal value, k_o, for steep than for shallow payoff matrices. Thus, PMM affects the steepness of the objective reward function and thus should affect the optimality of reward-maximizing decision criterion placement with steeper objective reward functions being associated with more nearly optimal decision criterion placement. Second, although PMA affects the maximum expected reward and thus the peak (or asymptote) of the objective reward function in Figure 9a, it does not change the shape or steepness of the objective reward function as shown in Figure 9b. Thus, reward-maximizing decision cri-

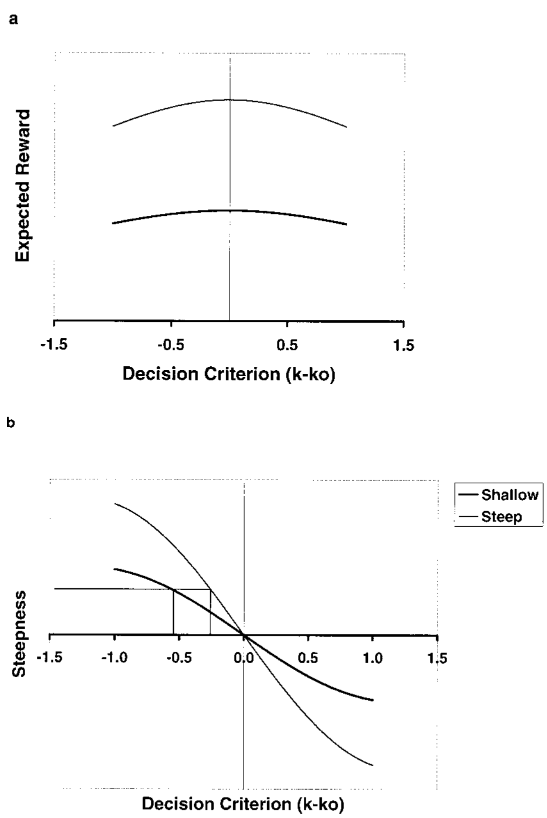

Fig. 9. Panel a: objective reward function for the shallow/no-cost and steep/no-cost conditions from Maddox, Dodd, and Bohil (2001). Panel b: steepness of the objective reward functions from Panel a.

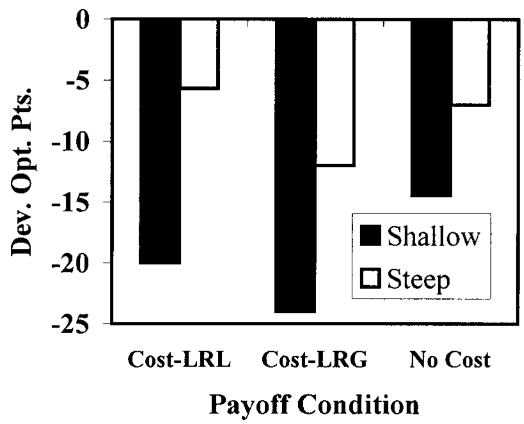

Fig. 10. Deviation from optimal points for a representative observer from Maddox, Dodd, and Bohil (2001).

terion placement should be unaffected by PMA (although the weight placed on accuracy in COBRA might). Finally, note that PMA and PMM do not affect the value of the optimal decision criterion.

Maddox et al. (2001) had 8 observers complete six 60-trial blocks in each of the eight 3:1 payoff conditions. Performance differed little across the two no-cost conditions, so these data were collapsed. Figure 10 plots the deviation from optimal points for a representative observer. As predicted from the flat-maxima hypothesis, performance was closer to optimal for the steep than for the shallow payoff matrices. Performance was also closer to optimal for the no-cost relative to the cost conditions.

Model-based analyses. The four base models of Figure 5 were applied simultaneously to the data from the eight payoff conditions in Table 1 separately by observer and block. A *hybrid* ($w_{cost\text{-}LRL}$; $w_{cost\text{-}LRG}$; $w_{no\ cost}$) *model* was also applied that estimated one accuracy weight from the two cost conditions that resulted in a long-run loss (i.e., the shallow and steep cost-LRL conditions), a second accuracy weight from the two cost conditions that resulted in a long-run gain (i.e., the shallow and steep cost-LRG conditions), and a third accuracy weight from the four nonnegative cost conditions (i.e., the shallow and steep no-cost [A] and no-cost [B] conditions). This model allowed us to determine whether the weight placed on accuracy differed across cost and no-cost conditions, as suggested by Maddox and Dodd (2001), and whether the weight placed on accuracy differed across long-run gain and long-run loss conditions. To determine whether observers were sensitive to the objective reward function steepness differences caused by PMM, the flat-maxima hypothesis was instantiated in two different ways. In one version, the PMM version, a single steepness parameter was estimated from the data that determined two distinct k_r values. One applied to the four steep payoff matrices, and the other applied to the four shallow payoff matrices. This version assumes that the observer is sensitive to steepness differences caused by payoff matrix multiplication, and is constrained to predict that the decision criterion will be closer to optimal for steep than for shallow payoff matrices as shown in Figure 9. In a second version, the no-PMM version, the single steepness parameter determined a single k_r value that applied to all eight conditions. In other words, this version assumed that the observer is not sensitive to steepness differences caused by payoff matrix multiplication, and instead behaves as if steepness of the objective reward function is equivalent across all eight conditions. Both versions of the flat-maxima model have one steepness parameter, and thus the fits of the two versions were compared directly. In fitting the two hybrid models, both versions of the flat-maxima hypothesis were tested.

To determine whether observers were sensitive to the steepness of the objective reward function as it relates to payoff matrix multiplication we compared the fits of the hybrid (w) and hybrid ($w_{\text{cost-LRL}}$; $w_{\text{cost-LRG}}$; $w_{\text{no cost}}$) models that assumed sensitivity to payoff matrix multiplication effects (i.e., the PMM versions) with those that assumed a lack of sensitivity (i.e., the no-PMM versions). For all 8 observers across all six blocks (except for 1 observer in Block 1), the version of the model that assumed sensitivity to payoff matrix multiplication, and thus more nearly optimal reward-maximizing decision criterion placement for steep than for shallow payoff matrices, provided the better account of the data. In addition, the model accounted for 90% to 92% of the responses in the data from each observer (91.5% of the responses in the data from the representative observer displayed in Figure 10).

Figure 11a displays the steepness, $w_{\text{cost-LRL}}$,

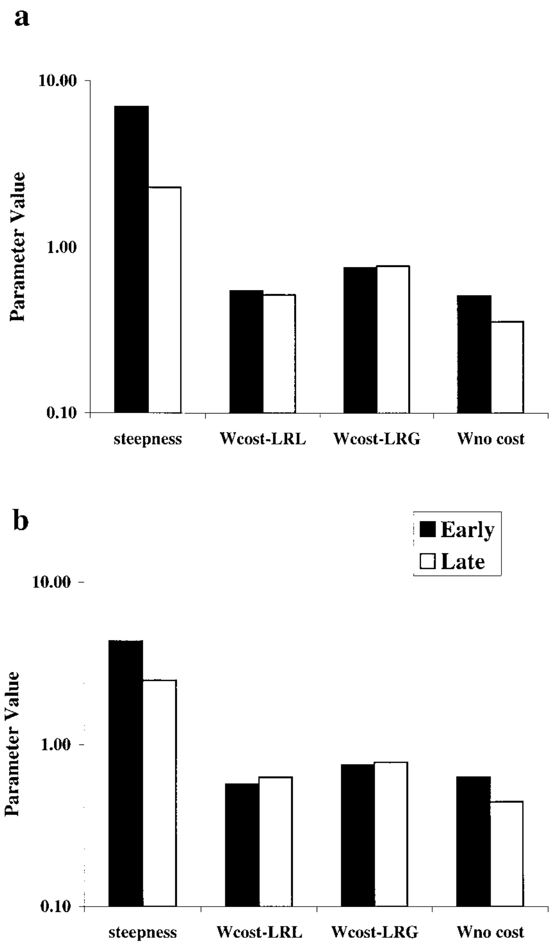

Fig. 11. Panel a: parameter values from the hybrid ($w_{\text{cost-LRL}}$; $w_{\text{cost-LRG}}$; $w_{\text{no cost}}$) model for a representative observer from Maddox, Dodd, and Bohil (2001). Panel b: parameter values from the hybrid ($w_{\text{cost-LRL}}$; $w_{\text{cost-LRG}}$; $w_{\text{no cost}}$) model averaged across fits of the model separately to each individual observer. The parameter values are plotted on a log scale and for the data early in learning (Blocks 1 and 2) and late in learning (Blocks 5 and 6).

$w_{\text{cost-LRG}}$, and $w_{\text{no cost}}$ values early and late in the session for a representative observer, and Figure 11b displays the same parameters averaged across fits to each individual observer's data. Several results are of interest. First, the steepness parameter declined from early to late in learning, replicating the finding from Maddox and Dodd (2001). Second, the accuracy weights remained relatively stable from early to late in the session in the payoff cost conditions, but declined from early to late in learning in the no-cost conditions. Third, more weight was placed on accuracy in the cost conditions for which a long-run gain was possible than in the cost conditions for which a long-run loss was expected. Although one might wish to speculate on the locus of this latter effect, this result requires replication before any strong claims can be made.

Decision Criterion Learning for Separate and Simultaneous Base-Rate and Payoff Manipulations

Until recently, nearly all studies of decision criterion learning focused on categorization problems in which base rates *or* payoffs were manipulated (however, see Healy & Kubovy, 1981). Given the fact that base rates and payoffs likely vary within the *same* natural-environment categorization problem, it is of interest to examine decision criterion learning in categorization problems that include simultaneous base-rate and payoff manipulations. Among other benefits, an examination of decision criterion learning under simultaneous base-rate/payoff conditions also allows a rigorous test of the *independence assumption* of the optimal classifier in human decision criterion learning. Recall from Equation 4 that the optimal classifier's decision criterion is determined from an independent combination of base-rate and payoff information. Encouraged by the successful application of the flat-maxima hypothesis to category discriminability and payoff matrix multiplication manipulations when base rates and payoffs were manipulated separately, Bohil and Maddox (2002)[3] generated predictions from the flat-maxima hypothesis for cases in which base rates and payoffs were manipulated simultaneously and to test those predictions empirically. Bohil and Maddox examined decision criterion learning for 16 observers in three 60-trial blocks in each of the following 10 base-rate/payoff conditions: 2:1B, 2:1P, 3:1B, 3:1P, 2:1B/2:1P, 3:1B/3:1P, 2:1B/3:1P, 3:1B/2:1P, 1:2B/3:1P, and 3:1B/1:2P, where B and P denote base rate and payoff, respectively. These conditions permitted a test of three a priori predictions from the flat-maxima hypothesis.

First, the flat-maxima hypothesis predicts

[3] Bohil, C. J., & Maddox, W. T. (2002). *A test of the optimal classifier's independence assumption in perceptual categorization.* Manuscript submitted for publication.

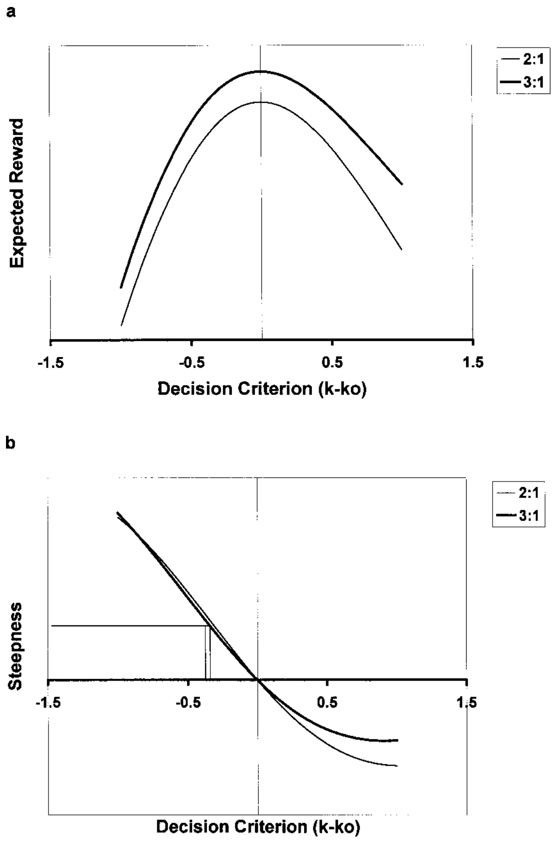

Fig. 12. Panel a: objective reward function for the 2:1 and 3:1 base-rate (or payoff) conditions from Bohil and Maddox (2002). Panel b: steepness of the objective reward functions from Panel a.

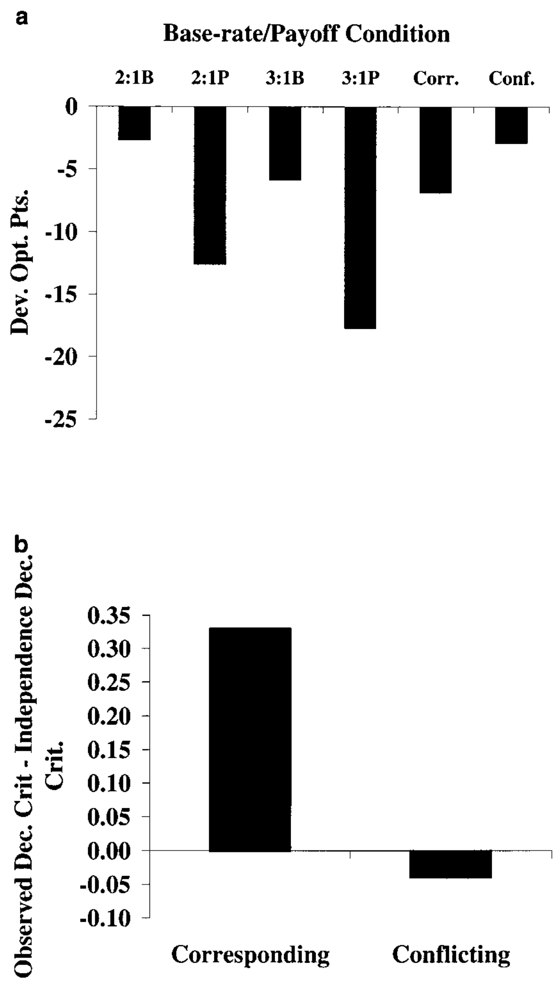

Fig. 13. Panel a: deviation from optimal points for a representative observer from Bohil and Maddox (2002). Panel b: observed decision criterion (from signal-detection theory) for the same representative observer minus the decision criterion predicted from the independence assumption.

superior performance in 2:1 over 3:1 conditions because the objective reward function is steeper in 2:1 conditions. Figure 12a displays the objective reward function for a 2:1 and a 3:1 base-rate (or payoff) condition. (The objective reward functions are identical for base-rate and payoff conditions with the same ratio.) Figure 12b plots the relation between the steepness for each objective reward function (i.e., the derivatives for each objective reward function) and $k - k_o$. Notice that for a fixed nonzero steepness, the decision criterion, k, is closer to the optimal value, k_o, for the 2:1 than for the 3:1 condition. Thus, the flat-maxima hypothesis predicts more optimal reward-maximizing decision criterion placement in 2:1 than in 3:1 conditions. Figure 13a displays the deviation from optimal points for the 2:1 and 3:1 base-rate and payoff conditions for a representative observer. These data show more nearly optimal performance in the 2:1 and 3:1 conditions as predicted from the flat-maxima hypothesis. Second, the flat-maxima hypothesis predicts that the observer's reward-maximizing decision criterion should be closer to the optimal value than that predicted from the optimal classifier's independence assumption in the *corresponding* simultaneous base-rate/payoff conditions. These are defined as simultaneous base-rate/payoff conditions for which the base rate and the payoff bias the observer toward the same categorization response, that is, the 2:1B/2:1P, 3:1B/3:1P, 2:1B/3:1P, and 3:1B/2:1P conditions. Third, the flat-maxima hypothesis predicts that the observer's reward-maximizing decision criterion should be farther from the optimal value than that predicted from the optimal classifier's independence assumption in the *conflicting* simultaneous base-rate/payoff conditions. These are defined as simultaneous base-rate/payoff conditions for which the base rate and the payoff bias the observer toward different categorization responses, that is, the 1:2B/3:1P and 3:1B/1:2P conditions.

The latter two predictions were generated as follows. First, we selected a single nonzero steepness value and derived the decision criterion from the associated objective reward function for the 2:1B, 2:1P, 3:1B, 3:1P, and the six simultaneous base-rate/payoff conditions. This yielded the decision criterion value for each of the six simultaneous base-rate/payoff conditions for that steepness. Second, for each of the six simultaneous base-rate/payoff conditions, appropriate 2:1B, 2:1P, 3:1B, and 3:1P decision criteria were combined (using Equation 4) to generate the decision criterion values for the six simultaneous base-rate/payoff conditions predicted from the independence assumption of the optimal classifier. Third, this process was repeated for a large number of nonzero steepness values. Note that, because the emphasis was on nonzero steepness values, the focus here was on cases in which a suboptimal decision criterion was utilized. Finally, for each of the six simultaneous base-rate/payoff conditions the log(β) value predicted from the flat-maxima hypothesis was plotted as a function of the log(β) value predicted from the independence assumption of the optimal classifier. These are depicted in Figure 14 by the solid curve. The broken line is included for comparative purposes, and denotes a situation in which the two hypotheses make identical predictions. Identical plots resulted for the 2:1B/3:1P and 3:1B/2:1P and for the 1:2B/3:1P and 3:1B/1:2P cases, so these were collapsed.

The most striking finding for the simultaneous base-rate/payoff conditions (Figures 14a through 14c) is that the decision criterion predicted from the flat-maxima hypothesis is always larger, and thus closer to the optimal value, than that predicted from the independence assumption of the optimal classifier. Thus, it follows that the flat-maxima hypothesis predicts that the decision criterion will be closer to the optimal value than that predicted from the independence assumption. Only when the 2:1B, 2:1P, 3:1B, and 3:1P decision criteria are optimal do the flat-maxima and independence-assumption decision criteria converge (not shown). Figure 14d plots the flat-maxima hypothesis and independence assumption predictions for the two conflicting simultaneous base-rate/payoff conditions. Here, the flat-maxima hypothesis predicts worse decision criterion learning than does the independence assumption of the optimal classifier. One advantage of these analyses is that specific violations of the independence assumption are predicted a priori, and, if these predictions are supported by the data, are well captured by the flat-maxima hypothesis without requiring any additional assumptions. Figure 13b displays the observed decision criterion value (derived from signal-detection theory; Green & Swets, 1967) minus the decision criterion value predicted from the independence assumption for the same representative observer whose data are displayed in Figure 13a. Notice that the value is positive for the corresponding conditions and is negative for the conflicting conditions. In other words, and in line with the flat-maxima hypothesis, decision criterion placement is closer to optimal than that predicted by the independence assumption for the corresponding conditions, but is farther from optimal for the conflicting conditions.

The four base models of Figure 5 were applied simultaneously to the data from each of the following 10 base-rate/payoff conditions: 2:1B, 2:1P, 3:1B, 3:1P, 2:1B/2:1P, 3:1B/3:1P, 2:1B/3:1P, 3:1B/2:1P, 1:2B/3:1P, and 3:1B/1:2P separately by observer and block. In addition, we applied a *hybrid* (w_P; w_{corr}; w_{conf}) *model* that estimated one accuracy weight for the two separate payoff conditions (i.e., 2:1P and 3:1P), a second accuracy weight for the four corresponding simultaneous base-rate/payoff conditions (i.e., 2:1B/2:1P, 3:1B/3:1P, 2:1B/3:1P, and 3:1B/2:1P), and a third accuracy weight for the two conflicting simultaneous base-rate/payoff conditions (i.e., 1:2B/3:1P and 3:1B/1:2P). The model accounted for 90% to 92% of the responses in the data from each observer (91.8% of the

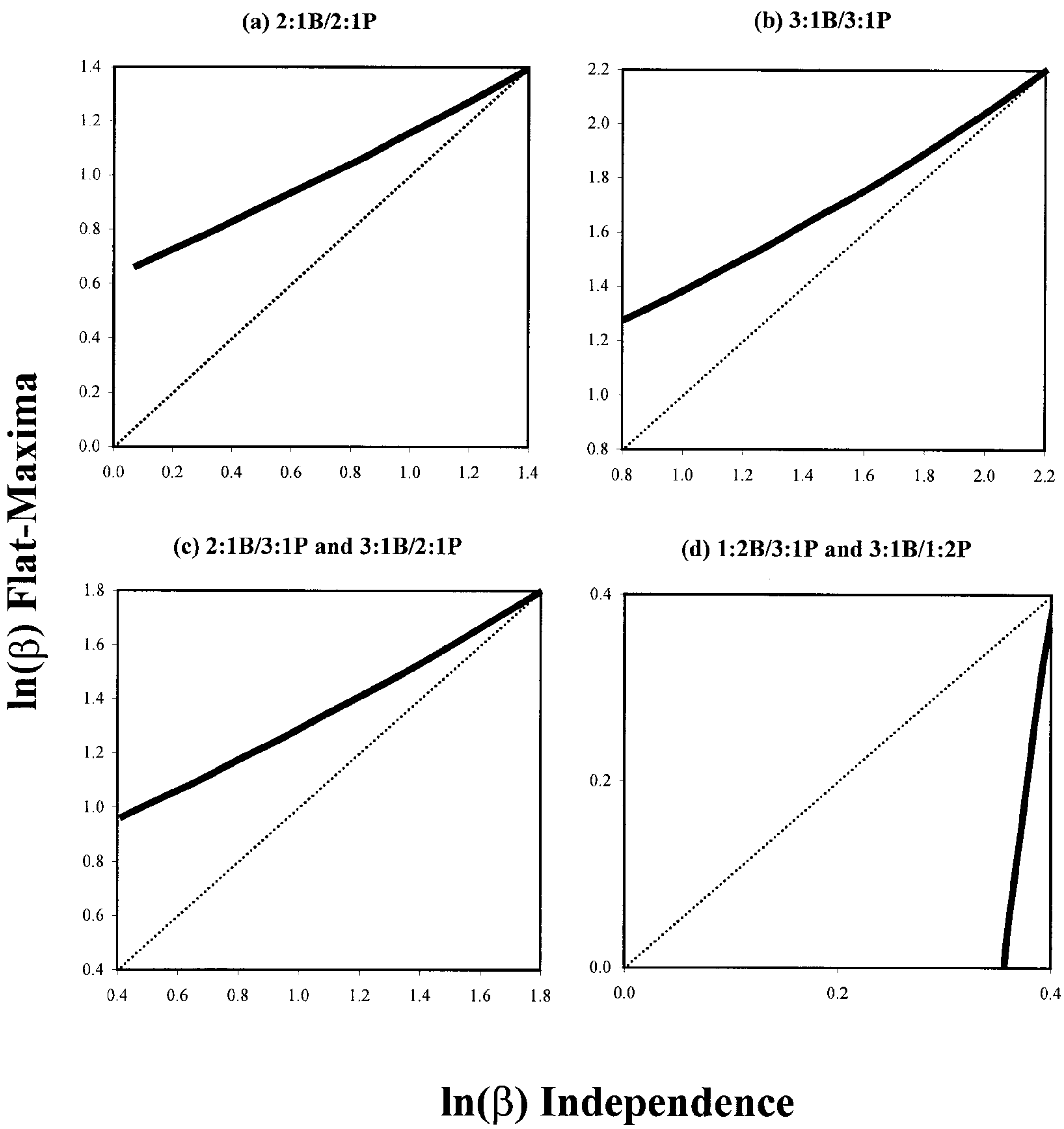

Fig. 14. Decision criterion, log(β), predicted from the flat-maxima hypothesis plotted against the decision criterion, log(β), predicted from the independence assumption of the optimal classifier for the six simultaneous base-rate/payoff conditions from Bohil and Maddox (2002). Panel a: 2:1B/2:1P condition. Panel b: 3:1B/3:1P condition. Panel c: 2:1B/3:1P and 3:1B/2:1P conditions. Panel d: 1:2B/3:1P and 3:1B/1:2P conditions. See text for details.

responses in the data from the representative observer displayed in Figure 13). Figure 15a displays the steepness, w_P, w_{corr}, and w_{conf} values for early and late learning for a representative observer, and Figure 15b displays the same parameters averaged across fits to each individual observer's data. The steepness values declined from early to late in the session, suggesting that the observer's reward maximizing decision criterion approached the optimal value. The accuracy weights suggested that the most weight was placed on accuracy in the separate payoff conditions, the least weight was placed on accuracy in the conflicting simultaneous base-rate/payoff conditions, and an intermediate amount of weight was

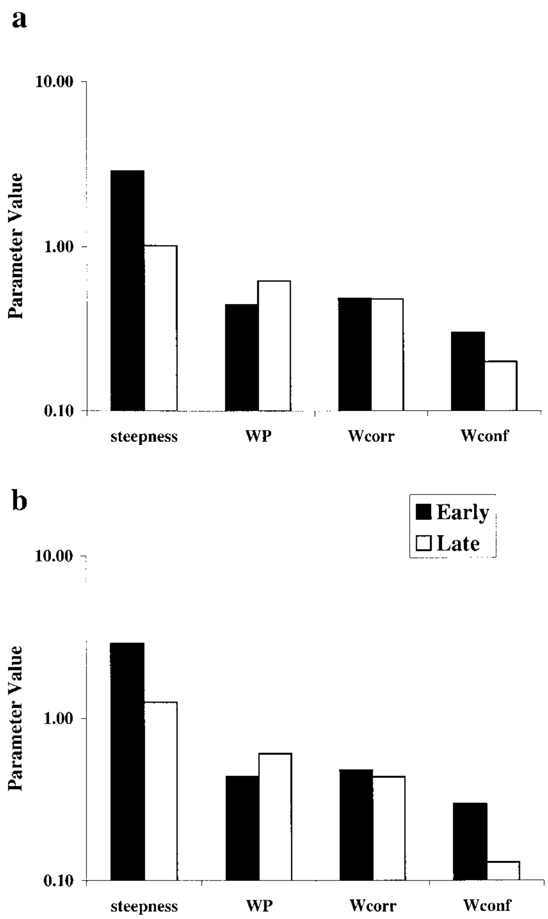

Fig. 15. Panel a: parameter values from the hybrid (w_P; w_{corr}; w_{conf}) model for a representative observer from Bohil and Maddox (2002). Panel b: parameter values from the hybrid (w_P; w_{corr}; w_{conf}) model averaged across fits of the model separately to each individual observer. The parameter values are plotted on a log scale for the data early in learning (Block 1) and late in learning (Blocks 2 and 3). P = payoff. Corr = corresponding. Conf = conflicting.

placed on accuracy in the corresponding simultaneous base-rate/payoff conditions. The fact that the weight placed on accuracy was low in the conflicting simultaneous base-rate/payoff conditions suggests that observers were more willing to sacrifice accuracy in the interest of reward maximization when the payoffs were biased toward one category while the base rates were biased toward the other category.

Model-based tests of the optimal classifier's independence assumption. Although the hybrid model provided a good account of the data and captured performance in the separate and simultaneous base-rate/payoff conditions, suggesting violations of the optimal classifier's independence assumption, the best test is to compare the hybrid model with a model that assumes an independent combination of base-rate and payoff information in simultaneous conditions. To achieve this goal, a variant of the hybrid (w) model was developed that assumed independence in the simultaneous base-rate/payoff conditions. In the resulting *hybrid (w) independence model,* we applied the hybrid model framework to determine the decision criteria in the separate base-rate and separate payoff conditions (i.e., 2:1B, 2:1P, 3:1B, and 3:1P), and then combined these independently (following Equations 2 and 4) to derive the decision criteria in the simultaneous base-rate/payoff conditions. (An "independence" variant of the hybrid (w_P; w_{corr}; w_{conf}) model cannot be developed because it contains accuracy weights associated specifically with the simultaneous base-rate/payoff conditions.) Note that the hybrid (w) and hybrid (w) independence models have the same number of parameters, so the fits can be compared directly.

In support of the flat-maxima hypothesis, for 72% of the data sets, the hybrid (w) model provided a better account of the data than did the hybrid (w) independence model. One advantage of this approach is that the models are identical in all respects except in whether they apply the flat-maxima hypothesis or independence assumption to the simultaneous base-rate/payoff conditions. In other words, the flat-maxima and independence models both assumed that the flat-maxima and COBRA hypotheses were valid in the separate base-rate/payoff conditions. Although this approach provides a statistically rigorous comparison of the flat-maxima and independence assumptions, one could argue that it biases the results in favor of the flat-maxima hypothesis, because the flat-maxima hypothesis is used to generate the decision criteria for the separate base-rate/payoff conditions. Another approach would be to develop a model that instantiated the independence assumption of the optimal classifier by freely estimating separate decision criterion values in the 2:1B, 2:1P, 3:1B, and 3:1P conditions and then combining these independently (following Equation 4). Although this approach lacks a rigorous underlying theory for determining the four separate base-rate/payoff condition decision criteria, it does appear to give the independence assumption the best chance of accounting for the data. The four-decision criterion parameter model was applied to the data and compared with the four-parameter hybrid (w_P; w_{corr}; w_{conf}) model. The hybrid (w_P; w_{corr}; w_{conf}) model provided a better account of the data for 3 of 16 observers in Trial Block 1 and for 10 of 16 observers in Blocks 2 and 3. So although the independence assumption was supported early in learning, the flat-maxima model was supported later in learning, after substantial experience had accrued. Taken together, these model-based comparisons suggest that the hybrid model, which embodies the flat-maxima and COBRA hypotheses, provides an excellent account of decision criterion learning across (a) different base-rate/payoff ratios, (b) separate base-rate/payoff manipulations, (c) corresponding simultaneous base-rate/payoff manipulations, and (d) conflicting simultaneous base-rate/payoff manipulations. The hybrid model accounts for the observed violations of the independence assumption without incorporating any additional assumptions.

Feedback Effects on Cost–Benefit Decision Criterion Learning

Most work on decision criterion learning uses trial-by-trial feedback based on the behavior of the objective classifier that obtains 100% accuracy. The feedback displays depicted in Figure 1 are based on the behavior of the objective classifier. The COBRA hypothesis postulates that decision criterion learning will be more suboptimal in payoff than in base-rate conditions because accuracy and reward can be maximized simultaneously in base-rate conditions, whereas some measure of accuracy must be sacrificed in payoff conditions to maximize reward. The optimal classifier is willing to sacrifice some measure of accuracy to maximize reward, whereas human observers seem to place importance on both goals resulting in behavior that does *not* maximize reward. Because of the importance of reward maximization in many natural-environment categorization problems, it is of interest to examine the effects of other types of feedback that might lead the observer to be more willing to sacrifice accuracy.

Maddox and Bohil (2001, Experiment 2) had 8 observers complete three 120-trial blocks to examine the effects of different types of corrective feedback on decision criterion learning in a 3:1 no-cost condition (the shallow/no-cost [A] condition from Table 1). The aim was to determine whether certain types of feedback might make observers more willing to sacrifice accuracy. Four experimental conditions were constructed from the factorial combination of two category discriminabilities ($d' = 1.0$ vs. 2.2) with two types of feedback (objective vs. optimal classifier). With *objective classifier feedback*, on each trial observers were told the number of points they earned on that trial (gain) and their cumulative total (total gain). They were also told the number of points that would have been earned had they responded with the objectively correct category label (potential gain) and the cumulative total had they been correct on every trial (total potential gain). With *optimal classifier feedback,* on every trial observers were told the number of points they earned on that trial (gain) and their cumulative total (total gain). They were also told the number of points that would have been earned by the optimal classifier and the cumulative total for the optimal classifier. Whereas the objective classifier will never respond incorrectly, on a certain proportion of trials the optimal classifier will respond incorrectly, but always in the interest of reward maximization. The hypothesis was that optimal classifier feedback might help observers learn to sacrifice accuracy to maximize reward.

Figure 16 plots the deviation from optimal points for the objective and optimal classifier feedback conditions at each level of d' for a representative observer. As predicted by the flat-maxima hypothesis, performance was closer to optimal for $d' = 2.2$ than for $d' = 1.0$. In addition, optimal classifier feedback led to superior performance compared with objective classifier feedback, suggesting that optimal classifier feedback helped the observer sacrifice accuracy in the interest of reward maximization. The four base models in Figure 5 were applied simultaneously to the data from all four conditions, separately by observer and block, and one additional model,

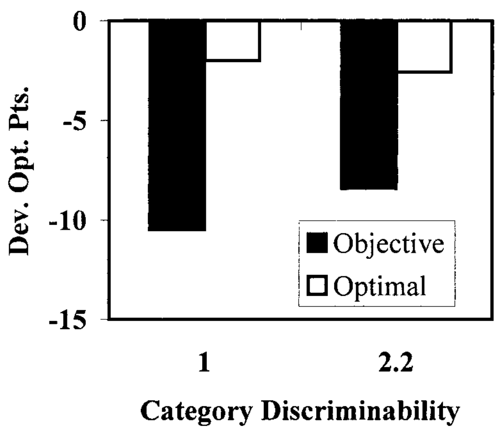

Fig. 16. Deviation from optimal points for a representative observer from Maddox and Bohil (2001).

the *hybrid* (w_{obj}; w_{opt}) *model* was applied to determine whether the weight placed on accuracy differed across feedback conditions. The prediction was that optimal classifier feedback would lead to less weight being placed on accuracy. The hybrid (w_{obj}; w_{opt}) model generally provided the most parsimonious account of the data, accounting for 88% to 92% of the responses in the data from each observer (91.9% of the responses in the data from the representative observer displayed in Figure 16). Figure 17a displays the steepness, w_{obj}, w_{opt} values for early and late learning for a representative observer, and Figure 17b displays the same parameters averaged across fits to each individual observer's data. Two results stand out. First, the steepness values declined from early to late in the session as in the three previous studies. Second, the weight placed on accuracy remained high and relatively stable in the objective classifier feedback condition (w_{obj}), and was lower and declined from early to late in the session for optimal classifier feedback (w_{opt}). This latter result suggests that observers were better able to sacrifice accuracy when feedback was based on the optimal classifier and continued to improve with experience.

CONCLUSION

This article outlines a recently developed unified theory of decision criterion learning and reviews studies that tested the generality of decision criterion learning under a wide range of experimental conditions. The theory assumes that two mechanisms are operative in decision criterion learning. One mechanism involves competition between reward and accuracy maximization. When both goals cannot be achieved simultaneously, the operative decision criterion falls somewhere between the accuracy- and reward-maximizing decision criteria. The second mechanism involves a flat-maxima hypothesis that assumes that the observer's estimate of the reward-maximizing decision criterion is determined from the steepness of the objective reward

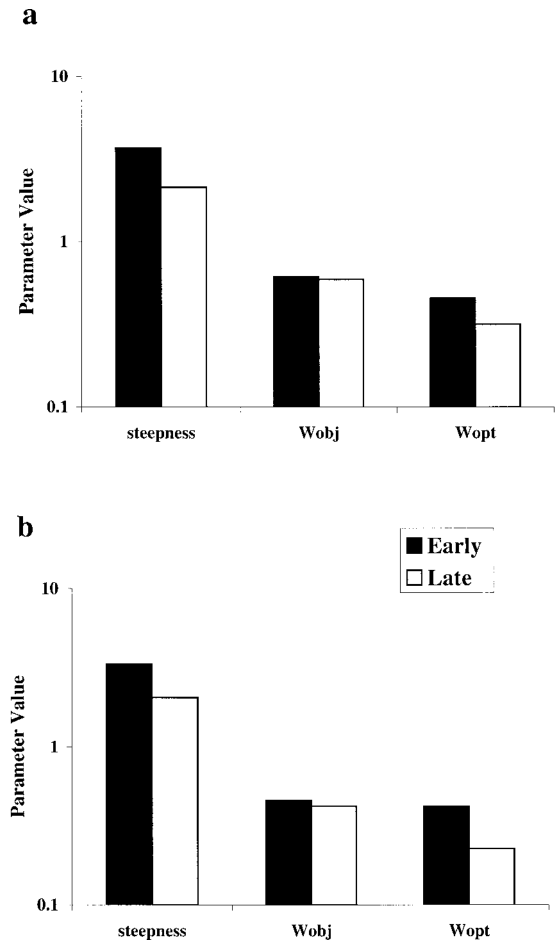

Fig. 17. Panel a: parameter values from the hybrid (w_{obj}; w_{opt}) model for a representative observer from Maddox and Bohil (2001). Panel b: parameter values from the hybrid (w_{obj}; w_{opt}) model averaged across fits of the model separately to each individual observer. The parameter values are plotted on a log scale for the data early in learning (Block 1) and late in learning (Blocks 2 and 3). Obj = objective classifier feedback. Opt = optimal classifier feedback.

function that relates expected reward to decision criterion placement.

Following the lead of most animal learning theorists, our participants completed several thousand trials and participated in all conditions of the experiment. This approach offers rigorous control over the participant's reinforcement history, and thus allows a better understanding of its effects on behavior. By examining a wide range of experimental conditions (i.e., reinforcement histories) within observer, we are able to directly compare behavioral profiles across reinforcement histories. This approach also focuses on individual behavior, as opposed to behavior aggregated over several participants. As shown in the 1950s (Estes, 1956) and reiterated more recently (Ashby et al., 1994; Maddox, 1999; Maddox & Ashby, 1998; Smith & Minda, 1998), averaging can alter the structure of data in such a way that the correct model of individual performance might provide a poor account of averaged performance, and even worse, an incorrect model of individual performance might provide an excellent account of averaged performance.

To date, this unified theory of decision criterion learning and several model-based instantiations has been applied to categorization problems that examine category discriminability, payoff matrix multiplication and addition effects, the optimal classifier's independence assumption, and different types of trial-by-trial feedback. For many of these manipulations, the flat-maxima and COBRA hypotheses make strong a priori predictions. In every case, a hybrid model framework that incorporated simultaneously the flat-maxima and COBRA hypotheses provided a good account of the data, and provided useful insights into the psychological processes involved in decision criterion learning. In particular, human observers appear to learn the optimal reward-maximizing decision criterion (or come very close to learning it) within 100 to 200 trials. In addition, the weight placed on accuracy is strongly affected by such factors as payoff matrix addition and the nature of the simultaneous base-rate/payoff condition (e.g., whether the base-rate and payoff ratios are corresponding or conflicting). Although all of this work was conducted with human subjects, our hope is that researchers interested in the optimality of nonhuman categorization behavior will find this approach useful and will incorporate it into their own research.

This body of research represents an important starting point, but much more work is needed to fully understand decision criterion learning. For example, all of this work was limited by the use of equal variance categories. Most natural categories likely do not have this property, so extensions to unequal category variances are in order. In addition, with the exception of a few early studies, all of this work used unidimensional stimuli. Although the dimensionality of the stimulus appears to have the largest effect on time needed to learn the category structures, extensions to multidimensional categories are still in order. Finally, all of the modeling work was conducted at the block-by-block level instead of the trial-by-trial level. The current models provide information about "average" performance within a block of trials, and by applying the models separately to each block of trials they provide information about learning. Other models have been proposed in the literature that model trial-by-trial changes in the decision criterion (e.g., Busemeyer & Myung's, 1992, hill-climbing model; Erev's, 1998, criterion reinforcement learning model; Wallsten & Gonzalez-Vallejo's, 1994, stochastic judgment model). An exciting avenue for future research will be to develop a trial-by-trial analogue of the hybrid model.

REFERENCES

Alsop, B. (1998). Receiver operating characteristics from nonhuman animals: Some implications and directions for research with humans. *Psychonomic Bulletin and Review, 5,* 239–252.

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., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105,* 442–481.

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of category learning. *Trends in Cognitive Sciences, 5,* 204–210.

Ashby, F. G., & Ell, S. W. (2002). Single versus multiple systems of learning and memory. In J. Wixted & H. Pashler (Eds.), *Stevens handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (3rd ed., pp. 655–692). New York: Wiley.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and 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., & Maddox, W. T. (1990). Integrating information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception and 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 and 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, judgement, and decision making* (pp. 251–301). New York: Academic Press.

Ashby, F. G., Maddox, W. T., & Bohil, C. J. (in press). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition.*

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., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Bohil, C. J., & Maddox, W. T. (2001). Category discriminability, base-rate, and payoff effects in perceptual categorization. *Perception & Psychophysics, 63,* 361–376.

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General, 121,* 177–194.

Davison, M., & Nevin, J. A. (1999). Stimuli, reinforcers, and behavior: An integration. *Journal of the Experimental Analysis of Behavior, 71,* 439–482.

Dusoir, A. (1980). Some evidence on additive learning models. *Perception and Psychophysics, 27,* 163–175.

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review, 105,* 280–298.

Erev, I., Gopher, D., Itkin, R., & Greenshpan, Y. (1995). Toward a generalization of signal detection theory to n-person games: The example of two person safety problem. *Journal of Mathematical Psychology, 39,* 360–375.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin, 53,* 134–140.

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001a). A possible role of the striatum in linear and nonlinear categorization rule learning: Evidence from patients with Huntington's disease. *Behavioral Neuroscience, 115,* 786–798.

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001b). Quantitative modeling of category learning in amnesic patients. *Journal of the International Neuropsychological Society, 7,* 1–19.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 234–257.

Green, D. M., & Swets, J. A. (1967). *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 and Memory, 7,* 344–354.

Herbranson, W. T., Fremouw, T., & Shimp, C. P. (1999). The randomization procedure in the study of categorization of multidimensional stimuli by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 113–135.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267–272.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243–266.

Herrnstein, R. J., & Heyman, G. M. (1979). Is matching compatible with reinforcement maximization on concurrent variable interval, variable ratio? *Journal of the Experimental Analysis of Behavior, 31,* 209–223.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399–1402.

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 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 and Psychophysics, 1,* 120–124.

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

Maddox, W. T. (1999). On the danger of averaging across observers when comparing decision bound and generalized context models of categorization. *Perception and Psychophysics, 61,* 354–374.

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

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 and Performance, 24, 301–321.
Maddox, W. T., & Bohil, C. J. (1998a). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1459–1482.
Maddox, W. T., & Bohil, C. J. (1998b). Overestimation of base-rate differences in complex perceptual categories. *Perception and Psychophysics, 60,* 575–592.
Maddox, W. T., & Bohil, C. J. (2000). Costs and benefits in perceptual categorization. *Memory & Cognition, 28,* 597–615.
Maddox, W. T., & Bohil, C. J. (2001). Feedback effects on cost-benefit learning in perceptual categorization. *Memory & Cognition, 29,* 598–615.
Maddox, W. T., & Dodd, J. L. (2001). On the relation between base-rate and cost-benefit learning in simulated medical diagnosis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1367–1384.
Maddox, W. T., & Filoteo, J. V. (2001). Striatal contribution to category learning: Quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's disease. *Journal of the International Neuropsychological Society, 7,* 710–727.
Mazur, J. E. (1988). Estimation of indifference points with an adjusting-delay procedure. *Journal of the Experimental Analysis of Behavior, 49,* 37–47.
Roth, A. E., & Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 3,* 3–24.
Shimp, C. P. (1966). Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior, 9,* 443–455.
Shimp, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review, 76,* 97–112.
Shimp, C. P. (1973). Probabilistic discrimination learning in the pigeon. *Journal of Experimental Psychology, 97,* 292–304.
Shimp, C. P., Long, K. A., & Fremouw, T. (1996). Intuitive statistical inference: Effects of sample size on categorizations of binomial samples. *Animal Learning & Behavior, 24,* 82–91.
Skinner, B. F. (1977). Why I am not a cognitive psychologist. *Behaviorism, 5,* 1–10.
Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1411–1436.
Staddon, J. E. R. (1992). Rationality, melioration, and law-of-effect models for choice. *Psychological Science, 3,* 136–141.
Staddon, J. E. R., & Ettinger, R. H. (1989). *Learning: An introduction to the principles of adaptive behavior*. San Diego, CA: Harcourt Brace Jovanovich.
Stephens, D. W., & Krebs, J. J. (1986). *Foraging theory.* Princeton, NJ: Princeton University Press.
Thomas, E. A. C. (1975). Criterion adjustment and probability matching. *Perception and Psychophysics, 18,* 158–162.
Thomas, E. A. C., & Legge, D. (1970). Probability matching as a basis for detection and recognition decisions. *Psychological Review, 77,* 65–72.
Ulehla, Z. J. (1966). Optimality of perceptual decision criteria. *Journal of Experimental Psychology, 71,* 564–569.
von Winterfeldt, D., & Edwards, W. (1982). Costs and payoffs in perceptual research. *Psychological Bulletin, 91,* 609–622.
Wallsten, T. S., & Gonzalez-Vallejo, C. (1994). Statement verification: A stochastic model of judgment and response. *Psychological Review, 101,* 490–504.
White, K. G., & Wixted, J. T. (1999). Psychophysics of remembering. *Journal of the Experimental Analysis of Behavior, 71,* 91–113.
Wickens, T. D. (1982). *Models for behavior: Stochastic processes in psychology.* San Francisco: Freeman.
Williams, B. A. (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, et al. (Eds.), *Stevens' handbook of experimental psychology: Vol. 1. Perception and motivation* (pp. 167–244). New York: Wiley.
Wishaw, I. Q. (1990). Time estimates contribute to food handling decisions by rats: Implications for neural control of hoarding. *Psycholobiology, 18,* 460–466.
Wishaw, I. Q., & Tomie, J. (1989). Food pellet size modifies the hoarding behavior of foraging rats. *Psychobiology, 17,* 93–101.
Wixted, J. T., & Gaitan, S. (in press). Cognitive theories as reinforcement history surrogates: The case of likelihood ratio models of human recognition memory. *Animal Learning & Behavior*.

Received January 28, 2002
Final acceptance June 16, 2002