

Optimal classifier feedback improves cost-benefit but not base-rate decision criterion learning in perceptual categorization

W. TODD MADDOX

University of Texas, Austin, Texas

and

COREY J. BOHIL

University of Illinois at Urbana-Champaign, Urbana, Illinois

Unequal payoffs engender separate reward- and accuracy-maximizing decision criteria; unequal base rates do not. When payoffs are unequal, observers place greater emphasis on accuracy than is optimal. This study compares objective classifier (the objectively correct response) with optimal classifier feedback (the optimal classifier's response) when payoffs or base rates are unequal. It provides a critical test of Maddox and Bohil's (1998) competition between reward and accuracy maximization (COBRA) hypothesis, comparing it with a competition between reward and probability matching (COBRM) and a competition between reward and equal response frequencies (COBRE) hypothesis. The COBRA prediction that optimal classifier feedback leads to better decision criterion learning relative to objective classifier feedback when payoffs are unequal, but not when base rates are unequal, was supported. Model-based analyses suggested that the weight placed on accuracy was reduced for optimal classifier feedback relative to objective classifier feedback. In addition, delayed feedback affected learning of the reward-maximizing decision criterion.

The need to categorize on the basis of uncertain information is ubiquitous in both personal and professional life. Each time we decide to follow or exceed the speed limit, to bring or not bring a jacket, to hire or not hire an individual, or to diagnose a patient, we are categorizing. All organisms categorize and must perform this task with some measure of success or they will die. In light of this fact, it is reasonable to suppose that, in many domains, human categorization performance is nearly optimal (Ashby & Maddox, 1998). Although optimality can be defined in a number of different ways, a common definition is performance that maximizes long-run reward (Green & Swets, 1966).

The optimal classifier is sensitive to information about category base rates (e.g., the prevalence of different diseases in the population) and the costs and benefits associated with correct and incorrect categorization (e.g., the benefit of correctly diagnosing a heart attack, and the cost of misdiagnosing) and uses this information to set a decision criterion that maximizes reward; values below

the criterion yield one categorization response, and values above the criterion yield another response. The optimal decision criterion is affected similarly by base-rate and cost-benefit manipulations (see Equation 2, below). For example, if the probability of observing an exemplar from Category A is three times that of observing an exemplar from Category B (a 3:1 base-rate condition), or if the benefit of a correct Category A response is three times the benefit of a correct Category B response (assuming no cost for an incorrect response; referred to as a 3:1 payoff condition), then the optimal decision criterion, $\beta_o = 3$, maximizes long-run reward.

Despite the survival value of optimal categorization, direct comparisons of human behavior with that of the optimal classifier suggest that humans rarely behave optimally. A robust finding in the categorization decision criterion learning literature is that observers use a sub-optimal decision criterion in both conditions, referred to as *conservative cutoff placement*, but that the magnitude of conservative cutoff placement is much larger when payoffs are manipulated than when base rates are manipulated, even when the optimal decision criterion is equated across conditions (Busemeyer & Myung, 1992; Erev, 1998; Healy & Kubovy, 1981; Maddox, 2002; for related work from the recognition memory literature, see Estes & Maddox, 1995; Heit, Brockdorff, & Lamberts, 2003). A thorough understanding of the mechanisms underlying these performance differences is of fundamental importance for at least two reasons. First, a better under-

This research was supported in part by National Institutes of Health Grant No. 5 R01MH59196 and by NIMH National Research Service Award No. MH14257 to the University of Illinois. We thank Kelli Hejl for help with data collection, and Evan Heit and three anonymous reviewers for helpful comments on an earlier version of this manuscript. Correspondence should be addressed to W. T. Maddox, Department of Psychology, 1 University Station A8000, University of Texas, Austin, TX 78712 (e-mail: maddox@psy.utexas.edu.)

standing of the cognitive processes involved in decision criterion learning under unequal base-rate and payoff conditions will help cognitive scientists develop and test computational models of performance in this ubiquitous, real-world task. A fruitful approach taken by a number of researchers is to use the optimal classifier as a benchmark against which to compare human performance and as a starting point for developing plausible computational models (Ashby & Maddox, 1993; Busemeyer & Myung, 1992; Erev, 1998; Maddox & Dodd, 2001). Second, a more thorough understanding will allow us to develop and test procedures that might lead to better decision criterion learning. Progress on these two general issues facing cognitive scientists provides the motivation for the present research.

Conservative Cutoff Placement, Probability Matching, and Accuracy Maximization

One popular explanation for the prevalence of conservative cutoff placement and the base-rate versus payoff decision criterion learning difference offered in the 1970s and 1980s was the generalized probability-matching hypothesis, which states that the participant chooses a decision criterion that yields response probabilities that match the base rates plus some constant based on the payoff matrix (Healy & Kubovy, 1981; see also Thomas & Legge, 1970). The generalized probability-matching hypothesis predicts conservative cutoff placement because the decision criterion that matches the response probabilities is generally suboptimal, and it predicts conservative cutoff placement of a greater magnitude in unequal payoff conditions because the base rates are equal and thus there is a bias toward the equal response frequency decision criterion. Specifically, in a 3:1 base-rate condition with category discriminability, $d' = 1.0$ (as in Healy & Kubovy, 1981), the decision criterion associated with 75% A and 25% B responses (i.e., probability matching) would be $\beta = 1.59$, whereas the optimal decision criterion is $\beta_o = 3$. Similarly, in a 3:1 payoff condition, the decision criterion associated with 50% A and 50% B responses would be $\beta = 1$, whereas the optimal decision criterion is $\beta_o = 3$. Healy and Kubovy found support for the generalized probability-matching hypothesis in a series of experiments that used the numerical decision task and a category $d' = 1.0$.

Maddox and Bohil (1998) provided some preliminary evidence against the generalized probability-matching hypothesis. Maddox and Bohil (1998) compared performance in a 3:1 base-rate and 3:1 payoff condition across three levels of d' (1.0, 1.7, and 2.2). Because the base rates and payoff matrices were fixed across d' conditions, the generalized probability-matching hypothesis predicts no effect of d' on decision criterion placement, but Maddox and Bohil (1998) found that the decision criterion was more nearly optimal for $d' = 1.7$ and 2.2 than for $d' = 1.0$. More rigorous tests have replicated this effect numerous times since (see Maddox, 2002, for a review).

Maddox and Bohil (1998) offered an alternative hypothesis. They suggested that a unitary goal, such as probability matching or reward maximization, may not underlie decision criterion learning; rather, participants may have multiple competing goals. Maddox and Bohil (1998) proposed a competition between reward and accuracy maximization (COBRA) hypothesis. COBRA postulates that observers attempt to maximize reward (consistent with instructions and monetary compensation contingencies) but also place importance on accuracy maximization.¹ When base rates are manipulated, the decision criterion that maximizes reward simultaneously maximizes accuracy and so there is no competition, leading to good decision criterion learning. On the other hand, when payoffs are manipulated, the decision criterion that maximizes reward does *not* simultaneously maximize accuracy. Because the observer must sacrifice some measure of accuracy to maximize reward when payoffs are manipulated, any weight placed on accuracy will lead to the use of a decision criterion that is more suboptimal than that used in the base-rate conditions.

Although COBRA predicts better decision criterion learning in base-rate than in payoff conditions, COBRA makes no assumptions about learning of the reward-maximizing decision criterion, and in isolation does not account for the fact that conservative cutoff placement is observed in both conditions. Maddox and Dodd (2001) developed a hybrid model of decision criterion learning that instantiates COBRA, but also assumes that the observer's estimate of the reward-maximizing decision criterion is determined from the objective reward function (von Winterfeld & Edwards, 1982). The quantitative details of the model will be outlined below, but for now a few brief comments are in order. The objective reward function plots long-run reward as a function of decision criterion placement (Figure 3, below). Steep objective reward functions, for which large changes in reward are associated with small changes in the decision criterion, are hypothesized to lead to better learning of the reward-maximizing decision criterion than flat objective reward functions. This is referred to as the *flat-maxima hypothesis*. One factor that influences the steepness of the objective reward function is category discriminability, with d' values near 2.2 yielding steep objective reward functions and smaller and larger values yielding flatter objective reward functions. Thus, the hybrid model predicts less conservative cutoff placement in base-rate and payoff conditions for steep objective reward functions (i.e., for d' values near 2.2) than for flat objective reward functions. Because many of the early studies used $d' = 1.0$ (e.g., Healy & Kubovy, 1981), which yields a flat objective reward function, the hybrid model would predict conservative cutoff placement for both base-rate and payoff conditions.

Despite the large body of evidence in support of COBRA, Maddox and Bohil (1998) may have been premature in rejecting probability matching in favor of accuracy maximization. It is worth stating explicitly that a

single-process model that assumes reward maximization, a single-process model that assumes accuracy maximization, and a single-process model that assumes probability matching all fail to account for the simultaneous effects of d' , base-rate, and payoff manipulations on decision criterion learning. So as a unitary process explanation, all three are inadequate. COBRA provides a good account of the data, but it is the “hybrid,” dual-process nature of the hypothesis that is the key to its success. Thus, a more appropriate test of the probability-matching hypothesis with COBRA would require that it be embedded within a dual-process framework. To achieve this goal, we develop below the competition between reward and probability matching (COBRM) hypothesis, which is identical to COBRA except that the accuracy-maximizing decision criterion is replaced with the probability-matching decision criterion.

Multiple-process models are difficult to test empirically using standard statistical procedures (e.g., analysis of variance [ANOVA]). Rigorous tests generally require that each observer participate in all experimental sessions, and that the data from all experimental conditions be modeled simultaneously. Using this approach, one can apply models that instantiate COBRA and COBRM and determine which provides the better account of the data. An even better approach is to combine quantitative modeling techniques with an experimental manipulation that provides a critical test of the two hypotheses. In the next section, we introduce an experimental manipulation of this sort that examines the effect of different training procedures on decision criterion learning.²

Objective Versus Optimal Classifier Feedback, Probability Matching, and Accuracy Maximization

As outlined earlier, when category base rates are manipulated, the decision criterion that maximizes reward simultaneously maximizes accuracy, and so COBRA predicts no competition. On the other hand, when payoffs are manipulated, the participant must sacrifice some measure of accuracy to maximize reward—a requirement that many participants appear unwilling to meet. Maddox and Bohil (2001) speculated that observers place importance on accuracy maximization in part because the most common type of feedback in decision criterion learning studies—objective classifier feedback—emphasizes accuracy. The top panel of Figure 1 depicts a hypothetical feedback display based on the objective classifier. Following a response, the observer is presented with information regarding the actual gain for that trial, and the potential gain had he/she responded with the correct category label. In this example, the observer generated an incorrect B response and earned 0 points, whereas a correct A response would have earned 3 points. (In our studies, we also include information regarding cumulative performance, i.e., the Total Points and the Potential Point Total). We refer to this as *objective clas-*

Objective Classifier	Categorization Response:	“B”
	Actual Gain:	0
	Potential Gain:	3
	Your Total Points:	28
	Potential Point Total:	56

Optimal Classifier	Categorization Response:	“B”
	Actual Gain:	0
	Optimal Classifier Gains:	0
	Your Total Points:	37
	Potential Point Total:	43

Figure 1. Hypothetical feedback displays for the objective classifier and optimal classifier feedback conditions.

sifier feedback, because the potential gain is always based on performance of the classifier that generates the objectively correct response on every trial, and thus is 100% accurate. Bohil and Maddox (2003; Maddox & Bohil, 2001) compared cost-benefit decision criterion learning under objective classifier feedback conditions, with feedback based on the optimal classifier. The bottom panel of Figure 1 depicts a hypothetical feedback display based on the optimal classifier. Following a response, the observer is presented with information regarding the actual gain for that trial, and the *optimal classifier’s* gain. In this example, the observer generated an incorrect B response and earned 0 points. Importantly, the optimal classifier also generated an incorrect B response and earned 0 points. Bohil and Maddox (2003) suggested that optimal classifier feedback might lead observers to sacrifice accuracy in order to maximize reward, resulting in better decision criterion learning.

Across a wide range of conditions, Bohil and Maddox (2003) found better decision criterion learning with optimal classifier as opposed to objective classifier feedback. They applied the hybrid model described above, which instantiates the flat-maxima and COBRA hypotheses, and found that the model provided a good account of the data by assuming that the weight placed on accuracy was greater for objective than for optimal classifier feedback. These data provide support for the COBRA hypothesis and suggest that optimal classifier feedback leads observers to place less weight on accuracy maximization, resulting in a more nearly optimal decision criterion. However, there are at least two alternative explanations that cannot be ruled out. One is the competition between reward and *probability matching* (COBRM) hypothesis outlined above. The longstanding interest in the matching–maximizing dichotomy makes this an important issue (e.g., Ashby & Maddox, 1993; Estes, 1976; Herrnstein, 1961, 1970; Herrnstein & Heyman, 1979;

Williams, 1988). A second alternative is referred to as the competition between reward and *equal response frequency* (COBRE) hypothesis (similar in spirit to the range frequency hypothesis developed by Parducci, 1965). The idea is that observers might attempt to maximize reward (as instructed) but might also be biased toward response patterns that yield equal “A” and “B” response frequencies. With respect to objective versus optimal classifier feedback, the assumption would be that optimal classifier feedback “releases” the observer from the bias toward probability matching (in COBRM) or from the bias toward equal response frequencies (in COBRE), resulting in more nearly optimal decision criterion placement.

Unfortunately, when costs and benefits are manipulated, all three hypotheses make identical predictions and thus are nonidentifiable. When costs and benefits are manipulated, base rates are equal and probability matching implies equal “A” and “B” response frequencies. Similarly, accuracy maximization implies equal “A” and “B” response frequencies. This equivalence of decision criterion values is depicted graphically in the top panel of Figure 2, which displays a hypothetical 3:1 payoff categorization problem with category discriminability, $d' = 1.0$. Notice that the accuracy-maximizing decision criterion (β_a), the decision criterion that satisfies probability matching (β_m), and the decision criterion that results in equal “A” and “B” response frequencies (β_e) all equal 1 (i.e., $\beta_a = \beta_m = \beta_e = 1$). Notice also that the optimal decision criterion $\beta_o = 3$.

To provide a critical test of these three hypotheses and to determine which best captures the effect of optimal versus objective classifier feedback on decision criterion learning, we need to examine an experimental situation for which each hypothesis makes a unique prediction. An unequal base-rate condition provides an excellent testing ground for at least three reasons. First, unequal base rates have the same effect on the optimal decision criterion as unequal costs and benefits, and so a base-rate and a cost-benefit condition that yield identical optimal decision bounds can be examined. Second, when base rates are unequal, the accuracy-maximizing, probability-matching, and equal response frequency decision criteria are different. This nonequivalence of decision criteria is depicted graphically in the bottom panel of Figure 2, which displays a hypothetical 3:1 base-rate categorization problem with $d' = 1.0$. Notice that the accuracy-maximizing decision criterion (β_a) is equivalent to the reward-maximizing decision criterion (β_o). The decision criterion that results in equal “A” and “B” response frequencies (β_e) is slightly less than 1.0, and the decision criterion that satisfies probability matching (β_m) lies intermediate between the equal response frequency and accuracy-maximizing decision criteria. Finally, because the accuracy- and reward-maximizing decision criteria are identical, COBRA predicts no effect of optimal classifier feedback relative to objective classifier feedback on base-rate learning since there is no competition. On

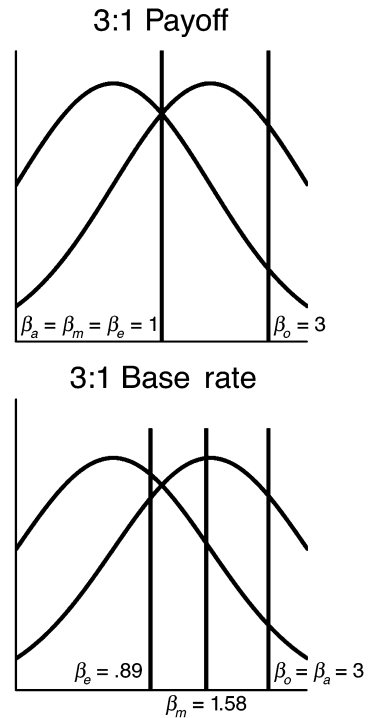


Figure 2. Hypothetical Category A and B distributions for $d' = 1.0$ under 3:1 payoff (top panel) or 3:1 base-rate (bottom panel) conditions. β_a , β_m , β_e , and β_o denote the decision criteria that maximize accuracy, satisfy probability matching, lead to equal “A” and “B” response frequencies, and maximize reward, respectively.

the other hand, if optimal classifier feedback releases the observer from a bias (either toward probability matching or equal response frequencies), then both the COBRM and COBRE hypotheses predict better base-rate learning with optimal classifier feedback than with objective classifier feedback.

This article reports the results from an experiment that examined the effects of optimal versus objective classifier feedback on decision criterion learning when payoffs or base rates and category discriminability ($d' = 1.0$ vs. 2.2) were manipulated. Finally, we revisited a delayed feedback manipulation initially investigated by Maddox and Bohil (2001; Experiment 1). In the immediate feedback condition, observers received information about their performance and that of either the optimal classifier or objective classifier on each trial. In the delayed feedback condition, feedback was not provided on every trial, but rather was provided on every fifth trial, and represented aggregate performance during those five trials. Maddox and Bohil (2001) speculated that delayed feedback might lead observers to focus less on accuracy, leading to better decision criterion learning. However, no effect of delay was observed. Thus, the present study consisted of 16 experimental conditions constructed from the factorial combination of two feedback condi-

tions (optimal vs. objective classifier feedback) with two base-rate/payoff conditions (3:1 payoff vs. 3:1 base rate), two category discriminabilities ($d' = 1.0$ vs. $d' = 2.2$), and two delay conditions (immediate feedback vs. delayed feedback). Each observer completed all 16 experimental conditions. A within-observers design was utilized to allow for model testing using the hybrid model, and variants of the hybrid model that instantiated the COBRM and COBRE hypotheses.

To summarize, the predictions are as follows. First, if optimal classifier feedback releases observers from a bias toward accuracy maximization, as suggested by COBRA, then we predict better decision criterion learning with optimal classifier feedback than with objective classifier feedback in the 3:1 payoff condition, but no decision criterion learning difference in the 3:1 base-rate condition. Second, if optimal classifier feedback releases observers from a bias toward probability matching or a bias toward equal response frequencies, as suggested by COBRM and COBRE, respectively, then we predict better decision criterion learning with optimal classifier feedback than with objective classifier feedback in both the 3:1 payoff condition *and* the 3:1 base-rate conditions. This follows because one cannot probability match or respond with equal frequencies in the 3:1 payoff and 3:1 base-rate conditions while simultaneously maximizing reward. Third, if delayed feedback has an effect similar to that of optimal classifier feedback, then the predictions outlined above will also hold for delayed relative to immediate feedback. Finally, we predict an interaction between the nature of the feedback (optimal vs. objective or delayed vs. immediate) and category d' on decision criterion learning, with a larger effect of feedback being predicted in the $d' = 1.0$ than in the $d' = 2.2$ conditions. This prediction holds because the accuracy sacrifice needed to maximize reward is larger for $d' = 1.0$ (8%) than for $d' = 2.2$ (3%), and the “A” to “B” response ratio assumed by the reward-maximizing decision criterion is larger for $d' = 1.0$ (.88 to .12) than for $d' = 2.2$ (.77 to .23).

Initial tests of these predictions are provided by examining trends in performance measures; specifically, point totals and decision criterion estimates from signal detection theory using ANOVA. These are adequate for many comparisons (e.g., feedback effects, d' effects, base-rate/payoff effects), but are not adequate for the more important comparisons between COBRA, COBRM, and COBRE. In particular, the differences between COBRM and COBRE are fairly subtle since both predict feedback effects on payoff and base-rate learning. These comparisons, and a detailed understanding of the psychological processes involved in decision criterion learning, are provided by applying a series of models to the data from all conditions simultaneously, but separately by observer and block. We now briefly review the optimal classifier and our modeling framework before turning to the experimental findings.

THE OPTIMAL CLASSIFIER AND DECISION BOUND THEORY

Optimal Classifier

Suppose the optimal classifier must determine whether a patient suffers from Disease A or B based on medical test X , whose outcomes for the diseases are normally distributed as depicted in Figure 2. The optimal classifier computes expected reward for each response and then chooses the response that maximizes expected reward. This requires comparing the likelihood ratio,

$$l_o(x) = f(x|B)/f(x|A), \quad (1)$$

with the *optimal decision criterion*,

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

and using the *optimal decision rule*:

$$\text{If } l_o(x) > \beta_o,$$

$$\text{then respond “B,” otherwise respond “A,”} \quad (3)$$

where $f(x|i)$ denotes the likelihood of test result x given disease category i , where $P(A)$ and $P(B)$ denote the Category A and Category B base-rate probabilities, respectively, and V_{aA} and V_{bB} denote the benefits associated with correct diagnoses. (The costs of incorrect diagnoses also affect the optimal decision criterion, but in the present study these were set equal to zero and thus drop out of the equation.)

Three points are of interest. First, when $P(A)V_{aA} = P(B)V_{bB}$, $\beta_o = 1$, the optimal classifier assigns the stimulus to the category with the highest likelihood and simultaneously maximizes reward, maximizes accuracy, probability matches, and generates equal response frequencies. Second, if the payoff for Disease A is three times the payoff for Disease B, a 3:1 payoff condition [i.e., if $V_{aA} = 3V_{bB}$ and $P(A) = P(B)$], $\beta_o = 3.0$. However, accuracy maximization, probability matching, and equal response frequencies must be sacrificed (see Figure 2, top panel). Finally, if the A base rate is three times the B base rate, a 3:1 base-rate condition [i.e., if $P(A) = 3P(B)$, and $V_{aA} = V_{bB}$], $\beta_o = 3.0$. In this case, accuracy and reward are simultaneously maximized, whereas probability matching and equal response frequencies must be sacrificed (see Figure 2, bottom panel).

Decision Bound Theory

There are at least two sources of noise in perceptual and cognitive systems—perceptual and criterial noise (Ashby & Lee, 1993; Ashby & Townsend, 1986)—that are not inherent in the optimal classifier decision rule (Equation 3). Perceptual noise refers to trial-by-trial variability in the perceptual information associated with repeated presentations of the same stimulus. With one perceptual dimension, the observer’s percept of stimulus i , on any trial, is given by $x_{pi} = x_i + e_p$, where x_i is the observer’s mean percept and e_p is a random variable

denoting perceptual noise (we assume that $\sigma_{pi} = \sigma_p$). (Notice that perceptual noise is different from the variability inherent in the category distributions.) Criterial noise refers to trial-by-trial variability in the placement of the decision criterion. With 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 denoting criterial noise (assumed to be univariate normally distributed). Decision bound theory assumes that the observer attempts to use the same *strategy* as the optimal classifier, but with less success because of the effects of perceptual and criterial noise (Ashby, 1992b). 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,

$$\text{If } I_o(x_{pi}) > \beta_o + e_c, \\ \text{then respond "B," otherwise respond "A."} \quad (4)$$

A THEORY OF DECISION CRITERION LEARNING AND A HYBRID MODEL FRAMEWORK

Maddox and Dodd (2001) developed a hybrid model of decision criterion learning that instantiates the flat-maxima and COBRA hypotheses.

Flat-Maxima Hypothesis

The flat-maxima hypothesis assumes that the observer adjusts the decision criterion (at least in part) on the basis of 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; Kubovy & Healy, 1977; Thomas, 1975; Thomas & Legge, 1970; von Winterfeldt & Edwards, 1982). The change in rate of reward is computed from the *objective reward function*, which plots objective expected reward on the y-axis and the decision criterion value on the x-axis (e.g., Busemeyer & Myung, 1992; Stevenson, Busemeyer, & Naylor, 1991; 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 repeated over a range of criterion values. The expected reward is then plotted as a function of decision criterion value. Figure 3a plots expected reward as a function of the deviation between a hypothetical observer's decision criterion $\ln(\beta)$ and the optimal decision criterion $\ln(\beta_o)$ standardized by category d' —referred to as $k - k_o = \ln(\beta)/d' - \ln(\beta_o)/d'$ —for category discriminability $d' = 1.0$ (solid line) and $d' = 2.2$ (broken line) and for the 3:1 payoff and 3:1 base-rate conditions. (The objective reward function is identical for 3:1 payoff and 3:1 base-rate conditions.) Notice that ex-

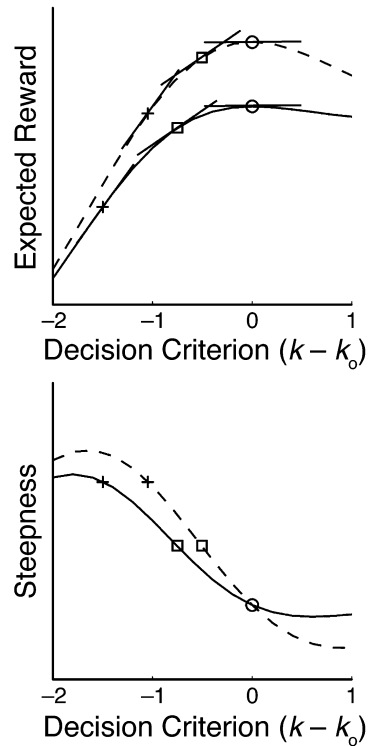


Figure 3. Objective reward functions for $d' = 1.0$ (solid line) and $d' = 2.2$ (broken line) for a 3:1 base-rate or 3:1 payoff condition (top panel). The tangent lines denoted by the plus sign have the same slope and thus reflect the same derivative or steepness. The same holds for the tangent lines denoted by the squares and circles. Steepness of the objective reward functions from the top panel, along with the “equal steepness” points denoted by the plus sign, square, and circle (bottom panel). Notice that the flat-maxima hypothesis predicts more nearly optimal decision criterion learning for $d' = 2.2$ than for $d' = 1.0$.

pected reward increases as the deviation from the optimal decision criterion decreases toward zero (i.e., the optimal decision criterion), and 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. Three derivatives are denoted by the plus sign, square, and circle and the associated tangent line in Figure 3a. The slope of each tangent line, corresponding to the derivative at that point, decreases as the deviation from the optimal decision criterion decreases (i.e., as we go from plus sign to square to circle). Notice also that the deviation from the optimal decision criterion for a fixed steepness value is smaller for $d' = 2.2$ than for $d' = 1.0$. 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. The flat-maxima hypothesis pre-

dicts that steeper objective reward functions ($d' = 2.2$) will lead to more nearly optimal decision criterion values than flatter objective reward functions ($d' = 1.0$). Because the flat-maxima hypothesis is based on the objective reward function, it applies only to learning of the reward-maximizing decision criterion. In the hybrid model, the observed decision criterion is a weighted average of the reward- and accuracy-maximizing decision criteria (i.e., the COBRA hypothesis).

COBRA

The second mechanism assumed to influence decision criterion placement is COBRA. Consider the 3:1 payoff condition depicted in Figure 4a. The reward-maximizing decision criterion, $k_{ro} = \ln(\beta_{ro})/d' = \ln(3)/d'$, is different from the accuracy-maximizing decision criterion, $k_{ao} = \ln(\beta_{ao})/d' = \ln(1)/d'$, and thus the observer cannot simultaneously maximize accuracy and reward. If an observer places importance or weight on reward and accuracy, then the resulting decision criterion will be intermediate between the reward- and accuracy-maximizing criteria. We instantiate this process with a simple weighting function,

$$k = wk_a + (1 - w)k_r, \tag{5}$$

where w ($0 \leq w \leq 1$) denotes the weight placed on accuracy. Equation 5 results in a single decision criterion, such as k_1 , that is intermediate between the accuracy- and reward-maximizing criteria. Figure 4d depicts a 3:1

base-rate condition. Here $k_{ro} = k_{ao}$ and so there is effectively no competition.

COBRM

COBRM postulates that observers attempt to maximize expected reward, but they also place importance on probability matching. Figure 4b depicts the 3:1 payoff case, and Figure 4e depicts the 3:1 base-rate case. The 3:1 payoff situation is exactly analogous to that from COBRA since the probability-matching decision criterion, k_m , is identical to k_a . However, in the 3:1 base-rate case, COBRM continues to predict a competition, whereas COBRA does not. As with COBRA, we instantiate this process with a simple weighting function,

$$k = wk_m + (1 - w)k_r, \tag{6}$$

where w ($0 \leq w \leq 1$) denotes the weight placed on probability matching. Equation 6 results in a single decision criterion, such as k_1 , that is intermediate between the probability-matching and reward-maximizing criteria.

COBRE

COBRE postulates that observers attempt to maximize expected reward, but they also place importance on generating equal numbers of ‘‘A’’ and ‘‘B’’ responses. Figure 4c depicts the 3:1 payoff case, and Figure 4f depicts the 3:1 base-rate case. The 3:1 payoff situation is analogous to that from COBRA and COBRM since the

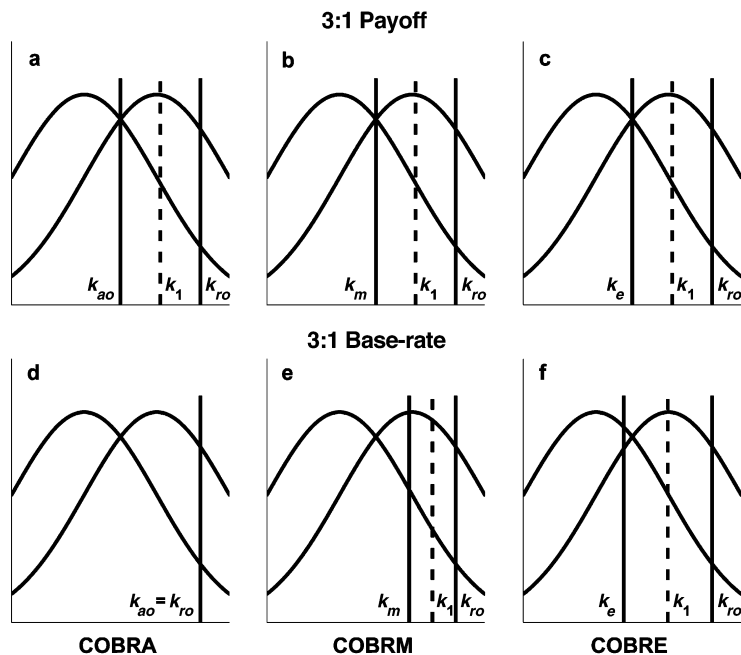


Figure 4. Hypothetical category distributions for $d' = 1.0$ under a 3:1 payoff condition (panels a–c) or 3:1 base-rate condition (panels d–f). Panels a and d instantiate the COBRA hypothesis, panels b and e instantiate the COBRM hypothesis, and panels c and f instantiate the COBRE hypothesis.

equal response frequency decision criterion, k_e , is identical to k_a and k_m . However, in the 3:1 base-rate case, COBRE (like COBRM) continues to predict a competition. We instantiate this process with a simple weighting function,

$$k = wk_e + (1 - w)k_r, \quad (7)$$

where w ($0 \leq w \leq 1$) denotes the weight placed on equal response frequencies. Equation 7 results in a single decision criterion, such as k_1 , that is intermediate between the equal response frequency and reward-maximizing criteria.

Framework for a Hybrid Model

The COBRA, COBRM, and COBRE versions of the hybrid model assume that the decision criterion used by the observer to maximize expected reward (k_r) is determined by the steepness of the objective reward function (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. Before each experimental condition, the observer is pretrained on the category structures in a baseline condition with equal base rates and equal payoffs (described in the Method section), which pretrains the accuracy-maximizing, probability-matching, and equal response frequency decision criteria since these are all equivalent under these conditions. COBRA, COBRM, and COBRE are instantiated in the hybrid model by estimating the w parameter in Equations 5, 6, and 7, respectively.

All of the models developed in this article are based on the decision bound model in Equation 4. Each model includes two “noise” parameters (one for $d' = 1.0$ and one for $d' = 2.2$) that represent the sum of perceptual and criterial noise (Ashby, 1992a; Maddox & Ashby, 1993). Each model assumes that the observer has accurate knowledge of the category structures [i.e., $l_o(x_{pi})$], which is reasonable since each observer completed a number of baseline trials and was required to meet a stringent performance criterion (see Method section). Finally, each model allows for a suboptimal decision criterion that is determined from the flat-maxima hypothesis along with COBRA, COBRM, or COBRE.

The nested structure of the models is presented in Figure 5, with each arrow pointing to a more general model and models at the same level having the same number of free parameters. The number of free parameters, including the two noise parameter described above, needed to fit a single block of data is presented in parentheses. (The details of the model-fitting procedure are outlined in the Results section.)

The simplest model was the hybrid model, which estimated a single steepness parameter from the data to determine k_r , and a single w parameter from the data. Three variants of the hybrid model, and all other models displayed in Figure 5, were examined. These included the COBRA, COBRM, and COBRE versions (see Equations

5–7). Four generalizations of the hybrid model were also developed. The hybrid(stp; w_{Optimal} ; $w_{\text{Objective}}$) model included a separate w parameter for optimal classifier feedback and objective classifier feedback. The hybrid(stp; $w_{\text{Immediate}}$; w_{Delay}) model included a separate w parameter for immediate and delayed feedback. The hybrid(stp; $w_{\text{Optimal/Immediate}}$; $w_{\text{Objective/Immediate}}$; $w_{\text{Optimal/Delay}}$; $w_{\text{Objective/Delay}}$) model also included a separate w parameter for each of the four feedback \times delay conditions. The final model was the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model. This model was developed after an initial examination of the results because, to anticipate, we observed a general performance decline across all conditions when feedback was delayed as opposed to immediate. We speculated that this general performance decline might be better captured by a deficit in reward-maximization decision criterion learning, and not to changes in the weight assigned to accuracy, probability matching, or equal response frequencies.

EXPERIMENT

The overriding goal of this experiment was to provide a critical comparison of the COBRA, COBRM, and COBRE hypotheses by examining the differential effects of objective versus optimal classifier feedback on decision criterion learning in 3:1 payoff and 3:1 base-rate conditions. Each observer completed 16 perceptual categorization tasks constructed from the factorial combination of two types of feedback (optimal and objective classifier) with two base-rate/payoff conditions (3:1 base rate and 3:1 payoff), 2 levels of d' (1.0 and 2.2), and two delayed feedback conditions (immediate or five-trial delayed). Each task consisted of three 120-trial blocks of training in which feedback was based on the optimal classifier or objective classifier, followed by a 120-trial test block, during which feedback was omitted. Table 1 displays the payoff matrix values, optimal points, optimal accuracy, and optimal decision criterion value for each experimental condition for a single block of trials.

Method

Observers. Six observers were recruited from the University of Texas community. All observers were tested and had 20/20 vision or vision corrected to 20/20. Each observer completed 16 sessions, each of which lasted approximately 60 min. Monetary compensation was based on the number of points accrued across the whole experiment. The data from one of the observers was excluded from all subsequent analyses because of poor performance. In many conditions, this observer perseverated on a single response, or responded randomly, rarely achieving performance above chance.

Stimuli and stimulus generation. The stimulus was a filled white rectangular bar (40 pixels wide) presented on the black background of a computer monitor. The bar rested upon a stationary base (60 pixels wide) that was centered on the screen, and bar height varied from trial to trial. There were two categories, A and B, whose members were sampled from separate univariate normal distributions. The sampled values determined the height of each presented bar stimulus. Category mean separation was 21 and 46 pixels for $d' = 1.0$ and $d' = 2.2$ conditions, respectively. The standard

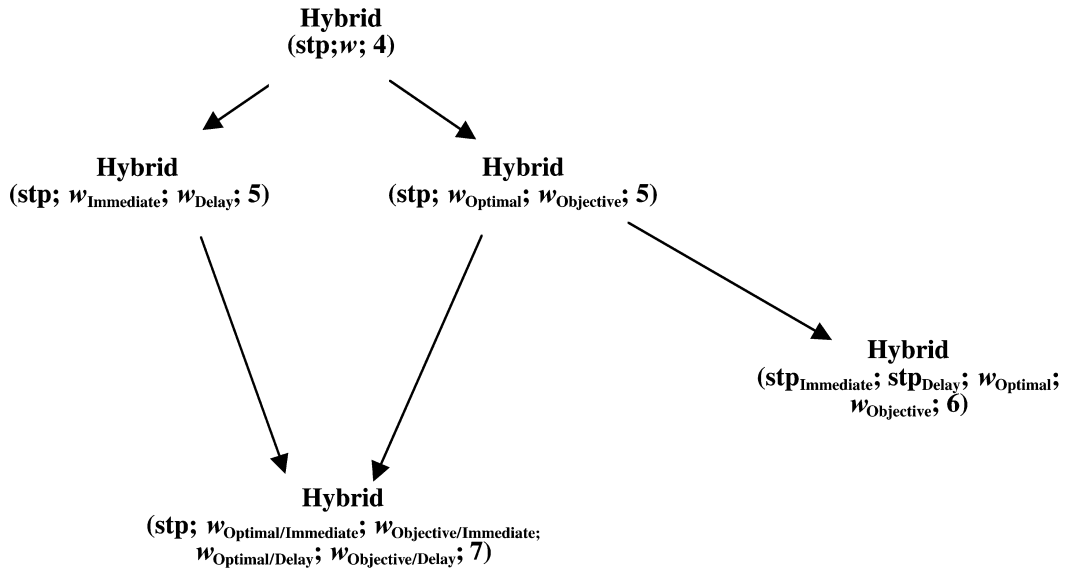


Figure 5. Nested relationship among the decision bound models applied simultaneously to the data from all experimental conditions. Each arrow points to a more general model. The number denotes the number of free parameters necessary to fit the model to a single block of trials.

deviation was 21 pixels for each category. Several random samples of size 60 were taken from each distribution, and the samples that best reflected the population means, standard deviations, and objective reward function were selected to yield a set of 120 unique stimuli for each level of d' .

Procedure. Prior to the first experimental session, observers were informed that they would be participating in a series of simulated medical-diagnosis tasks, and that on each trial of the experiment they would see a bar graph presented on the computer screen. They were told that the bar represented the result of a hypothetical medical test that was predictive of two possible diseases, and their job was to try to diagnose the patient on the basis of this test result (i.e., the height of the bar). The observers were told that each trial represented the test result for a new patient, and that they would earn a certain number of points for each trial depending on how they responded. They were instructed to try to maximize their point total over the course of the experiment, as this would determine their monetary compensation.

The order of presentation for the 16 experimental conditions was determined by Latin square, and observers completed one experimental condition during each daily session. To teach observers the category distributions prior to each experimental manipulation, as well as to minimize carryover effects, each experimental condition

was preceded by the completion of a minimum of 60 baseline trials, where category costs and benefits were unbiased (i.e., $V_{aA} = 2$, $V_{bB} = 2$). After observers had completed 60 baseline trials, performance was examined. If the observer reached an accuracy-based performance criterion (response accuracy not more than 2% below optimal), then those 60 trials were fit by two decision bound models (see Maddox & Bohil, 1998, for details). The *optimal decision criterion model* assumed that the observer used the optimal decision criterion (i.e., $\beta_o = 1$) in the presence of perceptual and criterial noise (Ashby, 1992b), whereas the *free decision criterion model* estimated the observer's decision criterion, along with perceptual and criterial noise, from the data. Because the optimal decision criterion model is a special case of the free decision criterion model, likelihood ratio tests were used to determine whether the extra flexibility of the free decision criterion model provided a significant improvement in fit. If the free decision criterion model *did not* provide a significant improvement in fit over the optimal decision criterion model, then the observer was allowed to begin the experimental condition. If the free decision criterion model *did* provide a significant improvement in fit, then the observer completed 10 additional trials, and the same accuracy- and model-based criteria were applied to the most recent 60 trials (i.e., Trials 11–70). This procedure continued until the observer reached the appropriate cri-

Table 1
Category Payoff Matrix Entries, Points, Accuracy, and Optimal Decision Criterion Value (Based on 120-Trial Blocks) for Each Experimental Condition

	Payoff Matrix Entries		$P(A)$	$P(B)$	Points	Accuracy	β_o
	VaA	VbB					
$d' = 1.0$, Baseline	2	2	.50	.50	166	69.2	1
$d' = 1.0$, 3:1 Payoff	3	1	.50	.50	186	61.0	3
$d' = 1.0$, 3:1 Base rate	2	2	.75	.25	186	77.8	3
$d' = 2.2$, Baseline	2	2	.50	.50	206	85.9	1
$d' = 2.2$, 3:1 Payoff	3	1	.50	.50	212	82.9	3
$d' = 2.2$, 3:1 Base rate	2	2	.75	.25	212	88.7	3

terion. Including these baseline trials and this fairly conservative accuracy- and model-based performance criterion ensured that each observer had accurate knowledge of the category structures before exposure to each experimental manipulation, and minimized the possibility of within-observer carryover effects from one experimental condition to the next. In addition, a different set of disease (i.e., category) labels was used in each experimental condition.

A typical trial proceeded as follows. The stimulus was presented on the screen and remained until a response was made. Observers were instructed to categorize each stimulus by pressing the appropriate button on the keyboard. Five lines of feedback that stayed on the screen until the observer pressed a key to move on to the next patient followed each response in the immediate feedback condition. Figure 1 presents hypothetical feedback displays for optimal classifier and objective classifier feedback under immediate feedback conditions. The top line of feedback indicated the disease possessed by the hypothetical patient (instead of the categorization response, which is depicted in Figure 1 for illustrative purposes). Fictitious disease names were used (e.g., "valinemia" or "brucellosis"), and a different pair of disease labels accompanied each experimental condition. The second line indicated the number of points gained for the given response. In objective feedback conditions, the third line displayed the potential gain for a (objectively) correct response on the trial. In other words, if the observer's response was correct, based on a priori category membership, then lines 2 and 3 of the feedback presented the same number of points. If the observer's response was incorrect, the third line showed what could have been earned had a correct response been given. In optimal classifier feedback conditions, however, the third line of feedback presented the number of points that the optimal classifier earned. In this case, if the observer made an (objectively) incorrect response (based on a priori category membership of the stimulus), but gave the correct response in relation to the optimal criterion, then both the observer and the optimal classifier would be incorrect for that trial, and lines 2 and 3 would present the same number of points gained or lost (Figure 1). The fourth line showed the number of points that the observer had accumulated to that point in the experimental condition, and the fifth line showed the number of points accrued by the objective or optimal classifier, depending on feedback condition. In the delayed feedback conditions, the top line of feedback was omitted. In addition, the remaining four lines of feedback denoted aggregate performance over the five trials. There was a 125-msec intertrial interval, during which the screen was blank, between removal of the feedback and presentation of the next stimulus. Observers were given a break every 60 trials.

Results and Theoretical Analysis

This section begins with an analysis of basic trends in the signal detection theory decision criterion estimates (Green & Swets, 1966) and point totals. We transformed each of these measures into a deviation from optimal score as follows:

deviation from optimal decision criterion

$$= k - k_o$$

$$= \ln(\beta)/d' - \ln(\beta_o)/d'$$

and

deviation from optimal points

$$= \frac{(\text{observed points} - 2 \text{ optimal points})}{\text{optimal points}}$$

These analyses will provide insights into the effects of d' , objective versus optimal classifier feedback, and de-

layed versus immediate feedback on payoff and base-rate decision criterion learning. They will also provide initial tests of the COBRA, COBRM, and COBRE hypotheses as they relate to the effects of objective versus optimal classifier feedback on base-rate and payoff learning. More rigorous tests will be introduced in the next section, devoted to the model-based analyses.

Performance Trends

A $2 d'$ (1.0 vs. 2.2) \times 2 feedback condition (objective vs. optimal classifier feedback) \times 2 base-rate/payoff condition (3:1 payoff vs. 3:1 base rate) \times 2 delay condition (immediate vs. delay) \times 4 block within-observers ANOVA was conducted in the deviation from optimal decision criterion and deviation from optimal point measures. The most important results are displayed graphically in Figure 6, with the left-hand column of plots displaying the $k - k_o$ results, and the right-hand column of plots displaying the deviation from optimal point results. The deviation from optimal decision criterion values for each participant in each of the 16 experimental conditions (averaged across blocks) are displayed in Table 2.

The most important finding was a significant interaction for both $k - k_o$ and points between objective versus optimal classifier feedback and base-rate/payoff condition [$k - k_o$, $F(1,4) = 64.28$, $MS_e = .023$, $p < .001$; points, $F(1,4) = 21.61$, $MS_e = .001$, $p < .01$]. This interaction is depicted in the top two panels of Figure 6. Post hoc analyses revealed a significant feedback effect in the payoff condition [$k - k_o$, $t(4) = 2.46$, $p = .07$; points, $t(4) = 3.83$, $p < .05$], suggesting better decision criterion learning with optimal relative to objective classifier feedback, that was nonsignificant in the base-rate condition [$k - k_o$, $t < 1$; points, $t(4) = 1.31$, $p > .05$]. These results support the COBRA hypothesis that optimal classifier feedback helps the observer focus more attention on reward maximization than on accuracy maximization. They provide less support for the COBRM and COBRE hypotheses, since both of these predict better payoff and base-rate learning with optimal classifier feedback. Even so, the more rigorous test of these hypotheses awaits the model-based analyses.

The interaction between d' and base-rate/payoff condition was significant for both measures [$k - k_o$, $F(1,4) = 10.57$, $MS_e = .121$, $p < .05$; points, $F(1,4) = 61.27$, $MS_e = .001$, $p < .001$] and is depicted in middle two panels of Figure 6. Post hoc analyses suggested that the effect of d' was greater in the payoff condition than in the base-rate condition. No other interactions were found to be significant. Several main effects were significant. In support of the flat-maxima hypothesis, the main effect of d' was significant [$k - k_o$, $F(1,4) = 110.58$, $MS_e = .166$, $p < .001$; points, $F(1,4) = 20.01$, $MS_e = .012$, $p < .05$], resulting in better performance for $d' = 2.2$ ($k - k_o = -.41$; points = $-.06$) than for $d' = 1.0$ ($k - k_o = -.88$; points = $-.12$). For both measures, the main effect of delay was significant [$k - k_o$, $F(1,4) = 16.47$, $MS_e = .901$, $p < .05$; points, $F(1,4) = 13.32$, $MS_e = .019$, $p < .05$], resulting in worse decision crite-

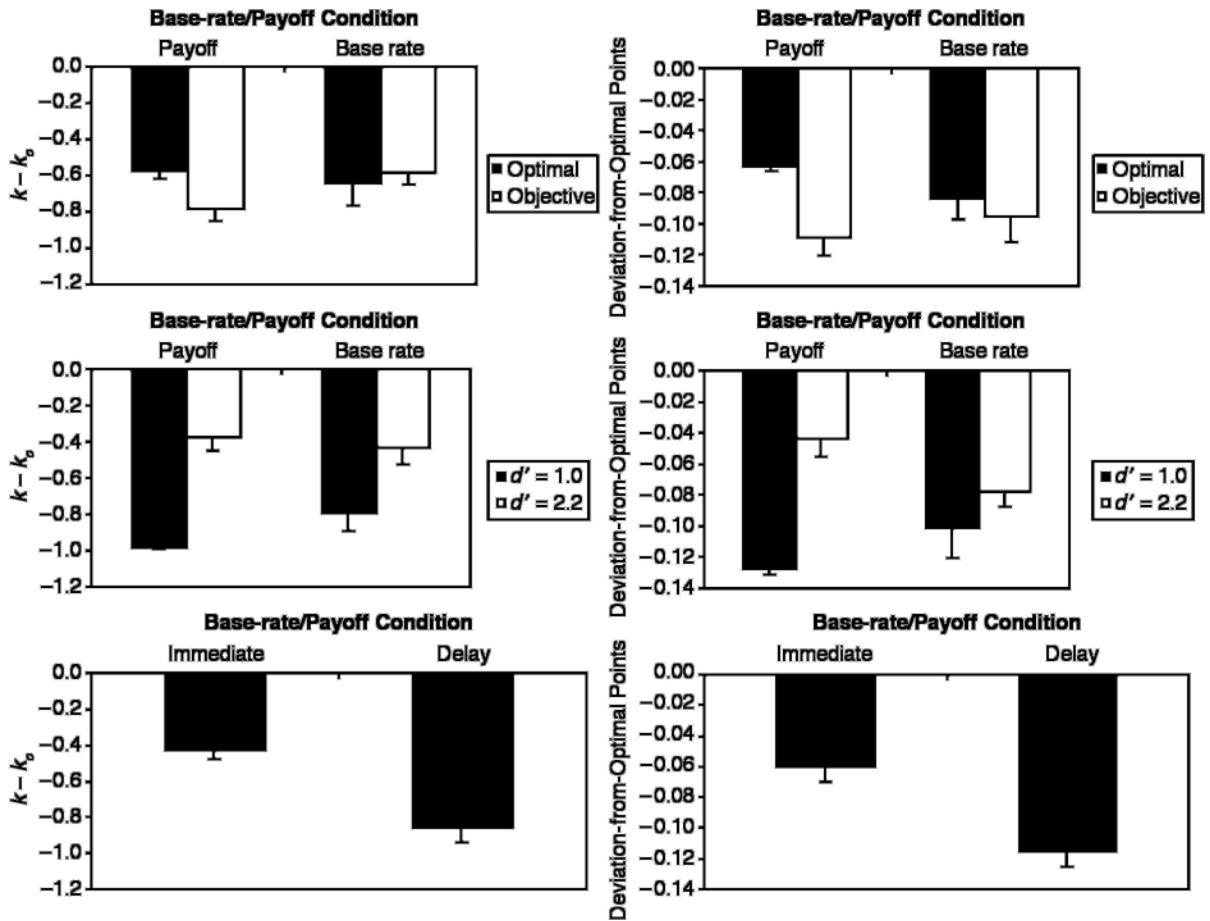


Figure 6. Average deviation from optimal decision criterion (left column) and average deviation from optimal points (right column) for the (top row) interaction between objective versus optimal classifier feedback and base-rate versus payoff condition, (middle row) interaction between d' and base-rate versus payoff condition, and (bottom row) delay effect. Standard error bars are included.

tion learning when feedback was delayed (see the bottom two panels of Figure 6). This factor did not interact with any other variables, suggesting that this deficit in learning was a general effect. The main effect of block was also significant [$k - k_o$, $F(3,12) = 9.82$, $MS_e = .082$, $p < .001$; points, $F(3,12) = 4.23$, $MS_e = .003$, $p < .05$], suggesting a gradual performance improvement with experience. The objective versus optimal classifier feedback main effect was significant for the point measure [$F(1,4) = 8.42$, $MS_e = .008$, $p < .05$], suggesting better decision criterion learning with optimal classifier ($-.07$) relative to objective classifier feedback ($-.10$), but was not significant for the $k - k_o$ measure ($F < 1$).

Before turning to the model-based analyses, we report on one additional analysis. For each participant in each d' and delayed feedback condition, we determined whether the decision criterion, k , and the point total were closer to optimal with optimal classifier feedback than with objective classifier feedback. Table 3 displays the proportion of times that each measure was closer to optimal

when feedback was based on the optimal classifier (collapsed across d' , delay, and observer) separately for the 3:1 payoff and 3:1 base-rate conditions and separately for each of the four blocks of trials. If COBRA is correct, and optimal classifier feedback helps the observer to focus more on reward maximization and less on accuracy maximization, then it should be the case that the decision criterion and point measures will be consistently closer to optimal with feedback based on the optimal classifier in the 3:1 payoff condition, but not in the 3:1 base-rate condition. On the other hand, both COBRM and COBRE predict that optimal classifier feedback should lead to more nearly optimal decision criteria and point totals in *both* the 3:1 payoff and 3:1 base-rate conditions. The results support COBRA. In the 3:1 payoff condition, the decision criterion measure was closer to optimal 75% of the time, and the point measure was closer to optimal 68% of the time, with feedback based on the optimal classifier. On the other hand, in the 3:1 base-rate condition, the decision criterion measure was

Table 2
Deviation From Optimal Decision Criterion ($k - k_o$) for Each of the 16 Experimental Conditions and 5 Observers (Averaged Across Blocks)

				Observer				
				1	2	3	4	5
$d' = 1.0$	Immediate	Objective	3:1 Base rate	-0.84	-0.58	-0.47	-0.86	-0.59
			3:1 Payoff	-1.25	-0.68	-0.27	-1.10	-0.69
		Optimal	3:1 Base rate	-0.47	-0.14	-0.41	-0.33	-0.75
			3:1 Payoff	-0.44	-0.29	-0.52	-0.88	-0.65
	Delay	Objective	3:1 Base rate	-1.34	-0.63	-1.16	-0.12	-1.50
				3:1 Payoff	-1.27	-2.14	-1.69	-1.04
Optimal		3:1 Base rate	-1.42	-0.98	-0.91	-0.90	-1.39	
			3:1 Payoff	-0.85	-0.70	-1.42	-0.96	-1.27
$d' = 2.2$	Immediate	Objective	3:1 Base rate	-0.60	-0.16	-0.06	-0.18	-0.01
				3:1 Payoff	-0.21	-0.33	-0.37	-0.46
		Optimal	3:1 Base rate	-0.49	-0.06	-0.41	-0.21	-0.10
				3:1 Payoff	-0.09	-0.30	0.00	-0.28
	Delay	Objective	3:1 Base rate	-0.46	-0.61	-0.16	-0.74	-0.60
				3:1 Payoff	0.28	-0.87	-0.65	-0.72
		Optimal	3:1 Base rate	-1.24	-0.23	-0.52	-0.23	-1.65
				3:1 Payoff	-0.48	-0.72	-0.22	-0.32

closer to optimal 48% of the time, and the point measure was closer to optimal 53% of the time with feedback based on the optimal classifier.

These analyses provide good initial support for the COBRA hypothesis that optimal classifier feedback will improve performance relative to objective classifier feedback in the 3:1 payoff, but not in the 3:1 base-rate condition. They also suggest that delayed feedback led to a general performance decrement that did not interact with any other factor (e.g., d' , base-rate/payoff condition, etc.). We turn now to the model-based analyses that more rigorously compare the COBRA, COBRM, and COBRE hypotheses, and determine whether delayed feedback led to a general or specific performance deficit.

Model-Based Analyses

Each of the five models displayed in Figure 5 was fit under the COBRA, COBRM, and COBRE assumptions, for a total of 15 models. Each model was applied simultaneously to the data from all 16 experimental conditions separately for each block and observer. Maximum likelihood procedures were used to estimate the parameters from each model, as they have a number of advantages over other estimation procedures (see Ashby, 1992b; Wickens, 1982, for details). Because the models were

applied separately to each block of trials, we could identify the model with the fewest free parameters that could not be improved upon (statistically) by a more general model, referred to as the *most parsimonious model*, for each observer in each block. However, our interest was in identifying the most parsimonious model overall for each observer. Since decision criterion shifts across trials are likely large early in learning, but are much smaller later in learning and during the test block, we decided to determine the most parsimonious model from the final training block and test block. Thus, the most parsimonious model was determined from the fit of each model summed over the last two blocks. (Using the cumulative fit across all four blocks did not change the pattern of results.) Even so, we do examine the parameter values separately by block to better characterize decision criterion changes with experience.

We took a two-pronged approach to the model-based analyses. First, we used a combination of likelihood ratio (G^2) tests (when the models were nested) and Akaike's AIC criterion tests (Akaike, 1974; for comparing non-nested models) to determine which of the five models in Figure 5 provided the most parsimonious account of the data when COBRA was assumed. We then repeated the process for the COBRM and COBRE models. The aim

Table 3
Proportion of Time That Optimal Classifier Feedback Led to More Nearly Optimal Performance Than Objective Classifier Feedback (Across d' and Delay Condition) for the $k - k_o$ and Deviation From Optimal Point Measures

		Block				Average
Measure	Condition	1	2	3	4	
$k - k_o$	3:1 Payoff	.75	.85	.70	.70	.75
	3:1 Base rate	.45	.50	.55	.40	.48
Dev. Opt. Points	3:1 Payoff	.65	.85	.60	.60	.68
	3:1 Base rate	.45	.55	.65	.45	.53

here was to determine whether there was some convergence across COBRA, COBRM, and COBRE variants regarding the steepness and weight parameter assumptions that provided the best fit to the data. Second, we compared the fits of the COBRA, COBRM, and COBRE models directly to determine which hypothesis best accounts for the effect of optimal versus objective classifier feedback on payoff and base-rate learning.

Best-fitting hybrid model assumptions for COBRA, COBRM, and COBRE. The first step was to determine the best fitting of the four hybrid models with one steepness parameter that assumed the COBRA hypothesis. On the basis of the likelihood ratio (G^2) tests of the maximum likelihood fit values (with $\alpha = .05$), the hybrid(stp; w) model provided the best account of the data from Observer 1, the hybrid(stp; w_{Optimal} ; $w_{\text{Objective}}$) model provided the best account of the data from Observer 4, and the hybrid(stp; $w_{\text{Optimal/Immediate}}$; $w_{\text{Objective/Immediate}}$; $w_{\text{Optimal/Delay}}$; $w_{\text{Objective/Delay}}$) model provided the best account of the data from Observers 2, 3, and 5. Next we compared the fit of the hybrid model with separate steepness parameters from the immediate and delayed feedback conditions [i.e., the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model] with that of the best fitting single steepness model using AIC. The results were clear. For Observer 1, the hybrid(stp; w) model continued to provide the best account of the data, but for the remaining 4 observers, the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model provided the best account of the data, suggesting that the delay manipulation did slow learning of the reward-maximizing decision criterion, whereas the objective versus optimal classifier manipulation affected the weight placed on accuracy maximization. It is worth emphasizing that the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model has fewer parameters than the hybrid(stp; $w_{\text{Optimal/Immediate}}$; $w_{\text{Objective/Immediate}}$; $w_{\text{Optimal/Delay}}$; $w_{\text{Objective/Delay}}$) model, and so it is not the case that the most general model provides a consistently superior account of the data. (In fact, for 4 of the 5 participants the hybrid[stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$] model with fewer free parameters is providing the better absolute fit of the data based on $-\ln L$.) This is important because it suggests that the models are capturing meaningful trends in the data and are not just overfitting the data.

The same approach was taken with the COBRM models, and we found that the hybrid(stp; $w_{\text{Optimal/Immediate}}$; $w_{\text{Objective/Immediate}}$; $w_{\text{Optimal/Delay}}$; $w_{\text{Objective/Delay}}$) model provided the best account of the data from Observers 3, and 5, and the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model provided the best account of the data from Observers 1, 2, and 4. Finally, assuming COBRE, we found that the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model provided the best account of the data from Observers 1, 2, 3, and 5, and the hybrid(stp; $w_{\text{Optimal/Immediate}}$; $w_{\text{Objective/Immediate}}$; $w_{\text{Optimal/Delay}}$; $w_{\text{Objective/Delay}}$) model provided the best account of the data from Observer 4.

Taken together, there was reasonable convergence across COBRA, COBRM, and COBRE models for the assumption that objective and optimal classifier feedback affects the weight placed on accuracy, probability matching, or equal response frequencies, and delayed feedback affects the speed with which the reward-maximizing decision criterion is learned. Next we directly compare the COBRA, COBRM, and COBRE variants of the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model.

Comparison of COBRA, COBRM, and COBRE variants of the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model. Because the three variants have the same number of parameters, we compared the fits directly. The results were clear. The COBRA variant provided the best account of the data from Observers 1, 2, 4, and 5, and the COBRM version provided the best account of the data from Observer 3. The absolute fit of the COBRA model was quite good, accounting for 83%–92% of the responses in the data from the 5 observers. This result converges with that from the ANOVAs in suggesting that the effect of optimal versus objective classifier feedback interacts with base-rate/payoff condition and that delayed feedback led to a general decline in decision criterion learning.

To determine how the observer's estimate of the reward-maximizing decision criterion was affected by the delayed feedback manipulation and to determine how these values changed across blocks, we examined the steepness parameters from the COBRA version of the hybrid(stp_{Immediate}; stp_{Delay}; w_{Optimal} ; $w_{\text{Objective}}$) model. These values are displayed in Figure 7a. To determine the magnitude of the weight placed on accuracy, how it was affected by the nature of the feedback, and how it changed across blocks, we examined the accuracy weight, w , parameters from the same model. These values are displayed in Figure 7b. A 2 delay type (immediate vs. delay) \times 4 block ANOVA revealed a main effect of delay type [$F(1,4) = 36.55, p < .01$], suggesting better reward-maximizing decision criterion learning in the immediate feedback condition, and a main effect of block [$F(3,12) = 9.50, p < .01$], suggesting learning with experience. The interaction was also significant [$F(3,12) = 3.89, p < .05$] and suggested a faster learning rate in the delayed feedback condition. Note, though, that initial reward-maximizing decision criterion learning in the immediate feedback condition is much better and then reaches ceiling in the second block of trials; thus it is likely not the case that delayed feedback leads to a faster rate of learning, but rather that immediate feedback quickly led to nearly optimal decision criterion learning. A 2 feedback type (objective classifier vs. optimal classifier) \times 4 blocks ANOVA was conducted on the accuracy weight values. Neither main effect was significant [feedback, $F(1,4) = 2.81, p > .05$; block, $F < 1$], but the interaction was [$F(3,12) = 3.70, p < .05$]. The most important finding is that the weight placed on accuracy was much lower, with optimal classifier feedback leading to better

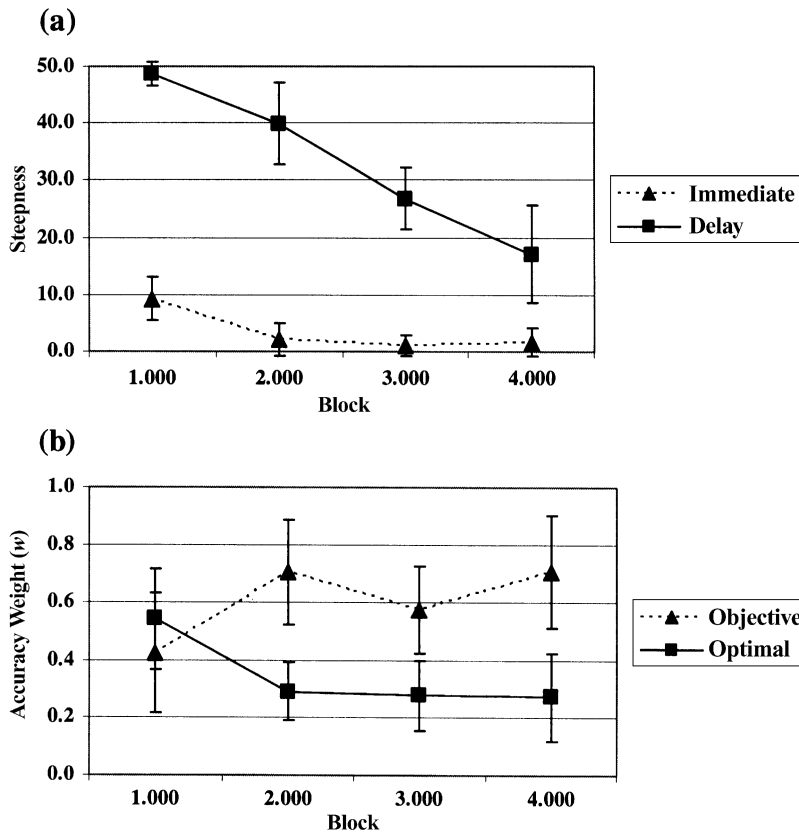


Figure 7. (a) Steepness values, and (b) accuracy weight, w , values from the hybrid($stp_{\text{Immediate}}$; stp_{Delay} ; w_{Optimal} ; $w_{\text{Objective}}$) model for the three training blocks and the test block averaged across observers. Standard error bars are included.

decision criterion learning. Interestingly, additional training beyond the first block had little effect on the accuracy weight in either feedback condition.

Before concluding, we compare the decision criterion values predicted from the COBRA, COBRM, and COBRE variants of the hybrid($stp_{\text{Immediate}}$; stp_{Delay} ; w_{Optimal} ; $w_{\text{Objective}}$) model with those observed in the experiment. This comparison might provide some insight into the successes and failures of the different model variants. Because the main result of interest was the interaction between objective versus optimal classifier feedback and base-rate versus payoff condition on decision criterion estimates, we compared the $k - k_o$ values from the three model variants with those generated from the data. These values (averaged across observers and blocks) are plotted in the four panels of Figure 8. Several comments are in order. First, notice that only the COBRA variant predicts the large effect of objective versus optimal classifier feedback on decision criterion learning in the payoff condition observed in the data, and predicts no effect of objective versus optimal classifier feedback on decision criterion learning in the base-rate condition (although the model predicts better decision criterion learning than

was observed in the data). Second, notice that the COBRM variant predicts a small effect of objective versus optimal classifier feedback on decision criterion learning in the payoff condition, but the magnitude of the effect is too small. In addition, as expected from an examination of Figure 4, the model predicts a similarly small effect of objective versus optimal classifier feedback on decision criterion learning in the base-rate condition that is not observed in the data. This variant also underpredicts the magnitude of the base-rate decision criterion learning deficit. Third, the COBRE model predicts no effect of objective versus optimal classifier feedback on decision criterion learning in either condition. This latter effect is somewhat surprising given the fact that the equal response frequency decision criterion is quite different from the reward-maximizing decision criterion in both conditions (see Figure 4). An examination of the w parameter estimates from the model provides an explanation for this result. These values are provided in Table 4 for the COBRE variant and for the COBRA and COBRM variants. Notice that the weight placed on equal response frequency was nearly zero for both types of feedback, in essence negating the predicted effect of objective versus

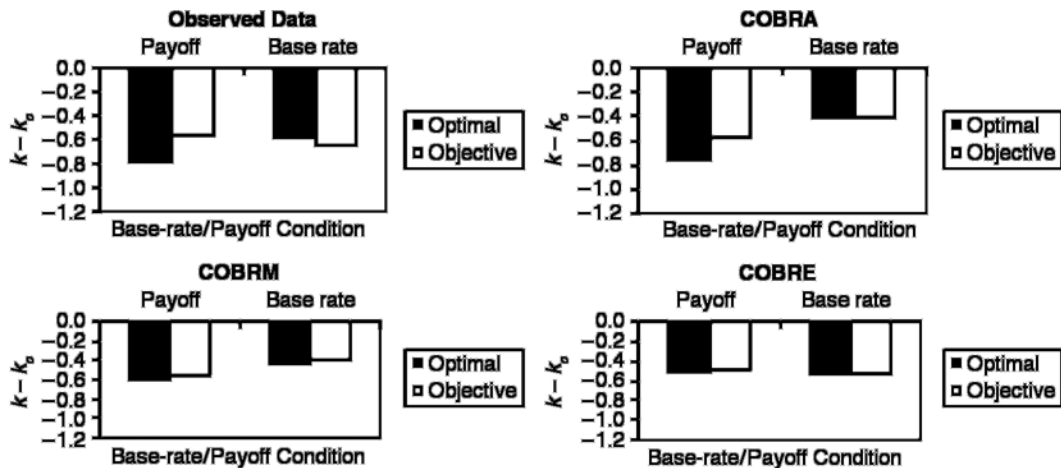


Figure 8. Deviation from optimal decision criterion for the interaction between objective versus optimal classifier feedback from the observed data, and the COBRA, COBRM, and COBRE versions of the hybrid($stP_{Immediate}$; stP_{Delay} ; $w_{Optimal}$; $w_{Objective}$) model averaged across observers and blocks.

optimal classifier feedback on decision criterion learning in base-rate and payoff conditions, leading to nearly equivalent predicted decision criterion values in all four conditions. This pattern most likely resulted because the model was faced with the task of predicting a pattern of results—optimal classifier feedback better than objective classifier feedback in payoff but not base-rate condition—that cannot be predicted by this model since the equal response frequency decision criterion was very different from the reward-maximizing decision criterion in both payoff and base-rate conditions. The COBRM model predicted a more reasonable pattern of results because it can predict a smaller effect of optimal classifier feedback in the base-rate condition than in the payoff condition because the probability-matching decision criterion is more similar to the reward-maximizing decision criterion in the base-rate than in the payoff condition. Even so, with reasonable weight placed on probability matching, the model is constrained to predict better decision criterion learning for optimal classifier feedback in both payoff and base-rate conditions, albeit a smaller effect for base rates.

GENERAL DISCUSSION

The overriding goal of the present research was (1) to obtain a more thorough understanding of the mechanisms underlying decision criterion learning differences observed in unequal base-rate and payoff conditions and

(2) to determine whether optimal classifier feedback, which has been found to improve decision criterion in payoff conditions, improves decision criterion learning in base-rate conditions. Single-process models such as those based on reward maximization, probability matching, or accuracy maximization have been offered in the literature to account for decision criterion learning under unequal base-rate and payoff conditions (e.g., Green & Swets, 1966; Healy & Kubovy, 1981; Maddox & Bohil, 1998). Each has had some limited success but provide inadequate accounts of the body of decision criterion learning data (see Maddox, 2002, for a review). One dual-process model has received some support. It assumes that decision criterion learning involves a competition between reward and accuracy maximization (COBRA). The present research compares the dual-process COBRA model with equivalent models that instantiate a competition between reward and probability matching (COBRM) and a competition between reward and equal response frequencies (COBRE). To achieve this goal, we compared goodness-of-fit across the three models, but more importantly, we included an experimental manipulation that provides a critical test of the three hypotheses by comparing decision criterion learning with objective versus optimal classifier feedback under unequal base-rate conditions.

Two previous studies (Bohil & Maddox, 2003; Maddox & Bohil, 2001) examined the effects of optimal versus objective classifier feedback on decision criterion learning when payoffs were unequal and found that optimal classifier feedback led to better decision criterion learning. Good accounts of the data were provided by COBRA and suggested that optimal classifier feedback reduced the weight placed on accuracy maximization, leading to more nearly optimal decision criterion learning. Unfortunately, COBRM and COBRE make identical predictions because the accuracy maximizing, probabil-

Table 4
Weight Parameter From the COBRA, COBRM, and COBRE Versions of the Hybrid($stP_{Immediate}$; stP_{Delay} ; $w_{Optimal}$; $w_{Objective}$) Model Averaged Across Block and Observer

	COBRA	COBRM	COBRE
Objective	.60	.52	.26
Optimal	.35	.33	.19

ity matching, and equal response frequency decision criteria are identical. However, the same nonidentifiability does not hold in unequal base-rate conditions. Because the reward- and accuracy-maximizing decision criteria are identical when base rates are unequal, COBRA predicts no effect of optimal classifier feedback on decision criterion learning, whereas both COBRM and COBRE predict an effect. In support of COBRA, optimal classifier feedback had a large effect on decision criterion learning in the payoff condition, but had no effect in the base-rate condition relative to objective classifier feedback. This result was supported by analyses of the point totals and signal detection theory decision criterion estimates, as well as by rigorous model-based analyses. The model parameters suggested that observers placed less weight on accuracy maximization when feedback was based on the optimal classifier. The COBRM model provided a reasonable account of the data, but underpredicted the performance advantage for optimal classifier feedback in the payoff condition and predicted an effect for the base-rate condition that was not observed in the data. The COBRE model provided the worst account of the data. In fact, the model predicted no effect of optimal classifier feedback on payoff or base-rate decision criterion learning and did not predict the robust finding of better decision criterion learning in base-rate relative to payoff conditions. Only COBRA predicted the overall pattern of decision criterion placement across objective versus optimal classifier feedback and base-rate versus payoff conditions.

Delayed Feedback

This study also examined the effects of delayed feedback on decision criterion learning. In the immediate feedback condition, observers received feedback on each trial. In the delayed feedback condition, observers received feedback on every fifth trial, and this feedback provided information about aggregate performance over those five trials. Maddox and Bohil (2001) speculated that delayed feedback might have an effect similar to that of optimal classifier feedback by leading the observer to focus less on accuracy maximization and more on reward maximization. This hypothesis was strongly rejected by the present data. In fact, delayed feedback led to consistently worse performance across *all* conditions than did immediate feedback. This effect was best modeled by assuming that delayed feedback slowed the learning of the reward-maximizing decision criterion and did not affect the weight placed on accuracy. Future research might examine shorter trial delays (e.g., every other trial) or temporal delays.

Training Implications

This work has implications for many real-world problems, in particular for training situations. In line with the results from Bohil and Maddox (2003), the present findings suggest that decision makers place importance on accuracy maximization and that they are often unwilling

to make the accuracy sacrifice necessary to maximize reward, and that this unwillingness is increased when the feedback emphasizes the accuracy of one's responding (objective classifier feedback). Beyond this, the present findings suggest that the nature of the feedback is not relevant to training when only base rates are manipulated since there is effectively no competition between reward and accuracy maximization. It is important to develop training procedures that reduce decision makers' emphasis on accuracy maximization when payoffs are being trained, but not when base rates are being trained. Training procedures based on the behavior of the optimal classifier, or possibly other desired classifiers, offer a promising approach.

Summary

In conclusion, the present study examined the effects of optimal versus objective classifier feedback on decision criterion learning when payoffs and base rates were manipulated, and provided a critical test of the COBRA hypothesis. The results suggest that the performance advantage observed for optimal classifier feedback relative to objective classifier feedback is specific only to unequal payoff conditions, where a sacrifice in accuracy maximization is necessary to maximize reward and does not result when base rates are unequal. This finding supports the COBRA hypothesis over COBRM or COBRE.

REFERENCES

- AKAIKE, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.
- 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., & LEE, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. In S. C. Masin (Ed.), *Foundations of perceptual theory* (pp. 369-399). Amsterdam: Elsevier.
- ASHBY, F. G., & MADDUX, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, *37*, 372-400.
- ASHBY, F. G., & MADDUX, 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., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154-179.
- BETTMAN, J. R., JOHNSON, E. J., LUCE, M. F., & PAYNE, J. W. (1993). Correlation, conflict, and choice. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 931-951.
- BOHIL, C. J., & MADDUX, W. T. (2003). On the generality of optimal versus objective classifier feedback effects on decision criterion learning in perceptual categorization. *Memory & Cognition*, *31*, 181-198.
- 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.
- DUSOIR, A. E. (1980). Some evidence on additive learning models. *Perception & Psychophysics*, *27*, 163-175.
- EREV, I. (1998). Signal detection by human observers: A cutoff rein-

- forcement learning model of categorization decisions under uncertainty. *Psychological Review*, **105**, 280-298.
- ERICKSON, M. A., & KRUSCHKE, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, **127**, 107-140.
- ESTES, W. K. (1976). The cognitive side of probability learning. *Psychological Review*, **83**, 37-64.
- ESTES, W. K., & MADDOX, W. T. (1995). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1075-1095.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- HEALY, A. F., & KUBOVY, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning & Memory*, **7**, 344-354.
- HEIT, E., BROCKDORFF, N., & LAMBERTS, K. (2003). Adaptive changes of response criterion in recognition memory. *Psychonomic Bulletin & Review*, **10**, 718-723.
- 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.
- 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.
- MADDOX, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior*, **28**, 1003-1018.
- MADDOX, W. T., & ASHBY, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, **53**, 49-70.
- MADDOX, W. T., & BOHIL, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1459-1482.
- MADDOX, W. T., & BOHIL, C. J. (2001). Feedback effects on cost-benefit learning in perceptual categorization. *Memory & Cognition*, **29**, 598-615.
- MADDOX, W. T., & BOHIL, C. J. (2004). Probability matching, accuracy maximization, and a test of the optimal classifier's independence assumption in perceptual categorization. *Perception & Psychophysics*, **66**, 104-118.
- 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, & Cognition*, **27**, 1367-1384.
- PARDUCCI, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, **72**, 407-418.
- PICKERING, A. D. (1997). New approaches to study of amnesic patients: What can a neurofunctional philosophy and neural network methods offer? *Memory*, **5**, 255-300.
- REBER, P. J., & SQUIRE, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory*, **1**, 217-229.
- RUSO, J. E., & DOSHER, B. A. (1983). Strategies for multi-attribute binary choice. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 676-696.
- SMITH, E. E., PATALANO, A., & JONIDES, J. (1998). Alternative strategies of categorization. *Cognition*, **65**, 167-196.
- STEVENSON, M. K., BUSEMEYER, J. R., & NAYLOR, J. C. (1991). Judgment and decision-making theory. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 283-374). Palo Alto, CA: Consulting Psychologists Press.
- THOMAS, E. A. C. (1975). Criterion adjustment and probability matching. *Perception & 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.
- VON WINTERFELDT, D., & EDWARDS, W. (1982). Costs and payoffs in perceptual research. *Psychological Bulletin*, **91**, 609-622.
- WICKENS, T. D. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco: Freeman.
- WILLIAMS, B. A. (1988). Reinforcement, choice, and response strength. In C. R. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' Handbook of experimental psychology: Vol. 2. Learning and cognition* (pp. 167-244). New York: Wiley.

NOTES

1. The idea of a competition between different goals or motivations is not new. Related ideas have been gaining support in the category-learning literature (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Erickson & Kruschke, 1998; Pickering, 1997; Reber & Squire, 1994; Smith, Patalano, & Jonides, 1998); this idea is similar in spirit to some of the effort-accuracy models of choice processing (Bettman, Johnson, Luce, & Payne, 1993; Russo & Doshier, 1983).

2. In a recent article that focused on tests of base-rate/payoff decision criterion learning independence, we provided a preliminary test of COBRA and COBRM and found mixed support for both (Maddox & Bohil, 2004). However, this article did not introduce an experimental manipulation that provided a critical test of these two hypotheses.

(Manuscript received November 18, 2003;
revision accepted for publication June 1, 2004.)