Signal Detection by Human Observers: A Cutoff Reinforcement Learning Model of Categorization Decisions Under Uncertainty

Ido Erev
Technion—Israel Institute of Technology

Previous experimental examinations of binary categorization decisions have documented robust behavioral regularities that cannot be predicted by signal detection theory (D. M. Green & J. A. Swets, 1966/1988). The present article reviews the known regularities and demonstrates that they can be accounted for by a minimal modification of signal detection theory: the replacement of the "ideal observer" cutoff placement rule with a cutoff reinforcement learning rule. This modification is derived from a cognitive game theoretic analysis (A. E. Roth & I. Erev, 1995). The modified model reproduces all 19 experimental regularities that have been considered. In all cases, it outperforms the original explanations. Some of these previous explanations are based on important concepts such as conservatism, probability matching, and "the gambler's fallacy" that receive new meanings given the current results. Implications for decision-making research and for applications of traditional signal detection theory are discussed.

Many common activities involve binary categorization decisions under uncertainty. While walking on campus, for example, students often try to distinguish between the individuals to whom they should say "hello" and the ones they had better ignore (uncertainty, in this case, arises from the limitations of individuals' memory and perceptual systems). The frequent performance of categorization decisions and the observation that they can have high survival value (as in the case of safety-related decisions) suggest that the cognitive processes that determine these decisions should be simple and adaptive. Thus, it could be hypothesized that one basic (simple and adaptive) model can be used to describe these processes within a wide set of situations.

The experimental literature provides mixed support for the simplicity and adaptivity hypothesis. The most impressive supportive evidence comes from applications of signal detection theory (SDT; see Green & Swets, 1966/1988). Under the assumption of adaptive (ideal) categorization decisions, this theory provides a good baseline for approximating and analyzing categorization decisions in a wide set of situations (e.g., see Davis & Parasuraman, 1982; Ferrel & McGoey, 1980; Macmillan & Creelman, 1991; Sorkin & Dai, 1994; Swennsen, Hessel, & Herman, 1977; Wallsten & Gonzalez-Vallejo, 1994). Experiments designed to provide a direct test of SDT have produced contradicting evidence. These studies, reviewed subsequently, reveal that experience does not always lead behavior toward the optimal response rule prescribed by SDT. Many robust violations of the optimal response rule do not disappear even when decision makers (DMs) gain the experience of hundreds of decision trials with immediate feedback.

The mixed evidence suggests that whereas some of the assumptions of SDT are likely to be robust, others have to be modified. The present article uses a cognitive game theoretic approach (Erev & Roth, in press; Roth & Erev, 1995) in an attempt to identify the robust assumptions and modify the weak assumptions in the search for a descriptive variant of SDT. As a means of facilitating generalizability, the current investigation starts with the modified signal detection model proposed by Erev, Gopher, Itkin, and Greenshpan (1995). Under this model, the choice process is abstracted by a cognitive interpretation of the law of effect (Thorndike, 1898). This model is more general than traditional SDT because it can be applied to binary categorization decisions in interactive settings (in addition to decisions under uncertainty).

The current article presents the logic behind the assumptions made by Erev et al. (1995) and then asks which assumptions must be added to their model to account for the robust behavioral regularities discovered in previous experimental studies of cate-

_________

1 As noted by Sperling and Dosher (1986), in the binary case, SDT is a corollary of subjective expected utility theory (Savage, 1954).

2 In a two-person categorization task, the payoffs of two observers are interdependent. For example, in a two-person safety dilemma (see Erev et al., 1995), an accident is avoided if at least one observer detects a warning signal. Thus, observers do not always know whether the outcome (e.g., no accident) reflects a safe state of nature or an accurate detection by their partner.
Signal detection theory (SDT) is a statistical theory of discrimination and classification based on the principles of signal detection and decision theory. It was developed by psychologist Percy Villee in the 1940s and was further refined by collaborators including Abraham Wald, Alexander B. Hayes, and Robert F. Mewhort.

SDT is used to analyze the performance of human observers in various tasks, such as visual and auditory signal detection, and is particularly useful in understanding the role of noise and signal in the decision-making process.

Three-Submodel Decomposition of SDT

In its basic form, SDT addresses a binary categorization task under uncertainty in which a DM (observer) is asked to decide how to label a stimulus (x) that may have come from one of two different sources, S1 (the noise distribution) or S2 (the signal distribution), with different probabilities. The theory is naturally decomposed into three cognitive game theoretic submodels, as follows.

- **State of Nature**: This refers to the underlying state of the world that is not directly observable, such as the presence or absence of a signal.
- **Response Space**: This refers to the possible responses of the observer, such as yes/no, go/no-go, or other binary choices.
- **Incentive and Information Structure**: This refers to the rewards and penalties associated with different outcomes, as well as the information available to the observer.

<table>
<thead>
<tr>
<th>State of Nature</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>R1</td>
</tr>
<tr>
<td>S2</td>
<td>R2</td>
</tr>
</tbody>
</table>

**Note.** S1 = noise distribution; S2 = signal distribution; R1 = noise response; R2 = signal response.

Incentive and information structure. SDT assumes that the information perceived by the observer can be summarized by the likelihood ratio of the stimulus given the two sources, that is, \( P(x|S2)/P(x|S1) \). Because the two distributions overlap, the observer cannot be 100% accurate. Rather, four contingencies are possible: The observer can correctly label the stimulus or make one of two possible errors.

**Strategy space.** The available cognitive strategies are assumed to be cutoff strategies. In the binary case, in which only two responses (R1 [noise] and R2 [signal]) are possible, each strategy can be summarized by the rule “response R2 if and only if the likelihood ratio \( P(x|S2)/P(x|S1) \) exceeds a certain cutoff.”

**Decision rule.** According to SDT’s ideal observer assumption, the observer is expected to select the cutoff that maximizes his or her expected utility. The optimal cutoff (the likelihood ratio \( \beta^* \)) can be calculated from the four contingencies and the prior probabilities.

Typical Experimental Setting and a Numerical Example

Most applications of SDT involve repeated decision tasks. Typically, the observer participates in multiple trials (often more than 100) given the same incentive structure and relevant stimuli sources. In each trial, the experimenter randomly chooses one of the two states (sources) in line with the relevant preset prior probabilities. Then the DM observes a stimulus drawn from the

---

3 Treisman (1987) studied signal detection without feedback; Ashby and Gott (1988) examined two-dimensional stimuli; Ward et al. (1988) evaluated the effect of sequential dependencies between stimuli; and Busemeyer and Myung (1992) addressed situations in which complex decision rules may be required.

4 Interestingly, this definition is consistent with the common use of the term in two distinct disciplines: cognitive psychology and classical game theory.

5 The utilities can also be scaled to set the lowest and highest values to 0 and 1, respectively.
selected source’s noisy distribution and has to choose among the response alternatives. In the studies considered here, feedback (typically points that convert to money at the conclusion of the experiment) was provided after each choice.

Whereas SDT does not distinguish between external noise (that characterizes the environment) and internal noise (the observer's ability), experimental studies typically consider only one of the two noise sources. Most studies focus on internal noise. In these studies, the experimenter has no control over the distribution of the perceived signal.

It is easier, however, to test (and explain) the theory when the perceived stimuli are controlled and observable by the experimenter. For that reason, much of the study of learning in probabilistic categorization tasks involves an external noise paradigm. In this paradigm (e.g., see Kubovy & Healy, 1977; Kubovy, Rapoport, & Tversky, 1971; Lee, 1963; Lee & Janke, 1964; Lee & Zentall, 1966), the experimenter controls the noise distributions. For instance, Kubovy and Healy (1977) presented numbers that represented heights of women (S1) and men (S2). In line with a common assumption, the heights in the two populations were normally distributed with equal variances and different means. Figure 1a shows the distributions studied by Kubovy and Healy (1977). Note that, in line with SDT's convention, the horizontal axis (c) measures distance in terms of the distributions’ standard deviation. The distance between the two distributions is referred to as d'. In Kubovy and Healy’s (1977) study, d' was 1.

SDT and Rationality

The understanding that the expected utility rule is often violated (e.g., see Kahneman & Tversky, 1979; Luce, 1959; Thaler, 1987; Tversky, 1969) has led many researchers (Busemeyer & Myung, 1992; Dorfman & Biderman, 1971; Kubovy & Healy, 1977; Lee, 1963; Schoeffler, 1965; Thomas, 1973) to hypothesize that violations of traditional SDT are likely to result from the weak descriptive power of the ideal observer cutoff placement assumption. Erev et al. (1995) built on Roth and Erev (1995) and examined the value of replacing this submodel with a reinforcement learning rule. We refer to this revised signal detection model as the cutoff reinforcement learning (CRL) signal detection model. Erev et al.'s results and the more recent findings of Gilat, Meyer, Erev, and Gopher (1997) and Gopher, Itkin, Erev, Meyer, and Armony (1995; summarized in Erev & Gopher, in press) appear to support this model. At least in the context of two-person signal detection tasks (studied in that research), the CRL signal detection model provides a good approximation of behavior. As noted earlier, the present article extends this analysis and asks which assumptions have to be added to this model to account for the main behavioral regularities observed in experimental studies of probabilistic categorization.

The Revised Decision (Learning) Rule

The basic idea behind Roth and Erev's (1995) learning model is a cognitive interpretation of the law of effect (Thorndike, 1898): the assumption that the probability that a certain strategy will be adopted increases when this strategy is positively reinforced. Similar models (with the exception of the cognitive interpretation) were suggested by Harley (1981; to describe animal learning processes), Bush and Mosteller (1955), and Luce (1959).

Erev et al. (1995) adapted the model to detection games under the assumption that the set of strategies available to the observers is a set of possible cutoffs. The adapted model’s basic assumptions are presented subsequently.

Finite Number of Uniformly and Symmetrically Distributed Cutoffs

According to the first assumption, the DM considers a finite number of m cutoffs. The m cutoffs are equally spaced along the interval \([c_{\text{min}}, c_{\text{max}}]\), where \(c_{\text{min}}\) and \(c_{\text{max}}\) are the two extreme cutoffs. The distribution is assumed to be symmetrical around \(c = 0\) (cf. Figure 1; i.e., \(c_{\text{min}} = -c_{\text{max}}\)), and the distance between two adjacent cutoffs is \(\Delta = 2(c_{\text{max}})/(m - 1)\). Thus, the first assumption is as follows.

**Assumption 1:** The DM considers a finite set of m cutoffs. The location of cutoff \(j\) \((1 \leq j \leq m)\) is \(c_j = c_{\text{min}} + \Delta(j - 1)\). Erev et al. (1995) set the two strategy-set parameters to \(m = 101\) and \(c_{\text{max}} = 5\). The x-axis in Figure 1 illustrates the assumed strategy set graphically. The locations of the 101 strategies are shown by tick marks.

**Initial Propensities**

The model assumes that the DM starts the experiment with a certain tendency (response strength or propensity) to choose each of the possible cutoffs. In line with Luce's (1959) distinction, the propensities are not probabilities (they can be larger than 1); as described later (see Assumption 4), however, they determine the choice probabilities.

**Assumption 2:** At time \(t = 1\) (before any experience), the DM has an initial propensity \(q_j(1) > 0\) to choose his or her jth cutoff. Erev et al. (1995) reduced the number of free initial propensity parameters to two by approximating the initial propensities distribution with a normal distribution having a mean at \(c = 0\). Thus, only two parameters, the initial distribution standard deviation \(\sigma_j\) and the area under the distribution, have to be set. As in Erev and Roth (1996), Erev et al. (1995) estimated the area as the expected absolute payoff from \(x(1)\) random decisions, where \(x(1)\) is a free parameter, and set the parameter \(x(1) = 3\). The standard deviation parameter was set to \(\sigma_j = 1.5\). The flatter distribution in Figure 1b illustrates this assumption.

---

6 The equal variance assumption is necessary to ensure one-to-one mapping between \(x\) and \(\beta\). Yet, the current model is likely to be robust to some violations of this assumption.

7 The general idea that the rationality assumption can be replaced by an adaptive learning process is becoming popular in economics (e.g., see Friedman, 1991; Fudenberg & Levine, 1996; Selten & Stoecker, 1986) and decision research (March, 1996).

8 The finite number is assumed to facilitate computer simulations. In principle, a continuous set of strategies is likely. Because the finite number can be very large, this computation assumption should not affect the predictions.
a. Conditional perceived stimuli distributions given the two states of nature

\[ d' = 1 \]

b. The generalization function relative to the initial propensities

\[ \text{The initial propensities} \]

Figure 1. Assumed strategy space and learning process. a: Signal detection theory's abstraction of the strategy space and the available information. Each tick mark represents a cutoff strategy. The two functions show the conditional probability distributions given the two possible states of nature (noise distribution \([S_1]\) and signal distribution \([S_2]\)). b: Initial propensities and generalization functions.

graphically (with Erev et al.'s parameters). The open bar above each strategy indicates the initial propensity to choose that strategy.

Reinforcement, Generalization, and Forgetting

The learning process is assumed to be a function of updating the propensities through reinforcement, generalization, and forgetting.

Assumption 3: If cutoff \( k \) was chosen at time \( t \) and the received payoff was \( v \), then the propensity to set cutoff \( j \) is updated by setting

\[ q_j(t + 1) = \max \{ [v(1 - \phi)q_j(t) + G_k[j, R(v, t))], \]

where \( v \) is a technical parameter that ensures that all of the propensities are positive, \( \phi \) is a recency (forgetting rate) parameter, \( G_k(\cdot, \cdot) \) is a generalization function, and \( R(\cdot, \cdot) \) is a reinforcement function.

As in Erev and Roth (1996), Erev et al. (1995) set the technical parameter to \( v = .0001 \) and the recency parameter to \( \phi = .001 \). The shape of the reinforcement and generalization functions was approximated on the basis of experimental results. Hermstein's (1961) demonstration of a linear relation between reinforcements and choice probabilities gave rise to a simple linear reinforcement function, \( R(v, t) = v - \rho(t) \), where \( \rho(t) \) is a reference point (in trial \( t \)); outcomes that are larger than \( \rho(t) \) are positive reinforcements, whereas outcomes that are smaller than \( \rho(t) \) are negative reinforcements. The observation that a reference point can move led Erev and Roth to assume the following contingent weighted average adjustment of the reference point:
where \( \rho(t) \) is the reference point at time \( t \) and \( w^+ \) and \( w^- \) are the weights by which positive and negative reinforcements affect the reference point. Following Erev and Roth, the initial reference point was set at \( \rho(1) = 0 \), and the weights were set at \( w^+ = .01 \) and \( w^- = .02 \).

Experimental investigations of generalization (e.g., Brown, Clark, & Stein, 1958) suggest a normal generalization distribution. To approximate a normal generalization distribution, Erev et al. (1995) assumed that

\[
G_s[j, R(v)] = R(v)\{F((c_j + c_j + 1)/2) - F((c_j + c_j - 1)/2)\},
\]

where \( F[.] \) is a cumulative normal distribution with mean \( c_j \) and standard deviation \( \sigma_p \). Erev et al. set \( \sigma_p = .025 \). A numerical example of a generalization function is presented graphically in Figure 1b. This example involves the generalization of a payoff of 2 cents for a choice of cutoff .3 in the first round (when the reference point is 0) under Erev et al.’s parameters and Kubovy and Healy’s (1977) task (in which the possible payoffs were +2 cents and -2 cents). The relative size of the area under the generalization and initial functions (1:3) reflects the fact that, in the current example, the reinforcement (area under the generalization function) is 2 and the area under the flat initial distribution is 6, \( \sigma(1)(\text{absolute payoff from random choice}) = 3 \times 2 \). Note that, as the DM gains experience, this ratio is expected to decrease (because the reference point moves toward the average payoffs, whereas the area under the propensities distributions increases as propensities are accumulated from round to round). Thus, the learning process is expected to display the “power law of practice” (Blackburn, 1936; Crossman, 1959).

Relative Propensities Sum

Following Luce (1959), the final assumption states a relative propensities sum choice rule.

**Assumption 4:** The probability that the observer sets strategy \( k \) at time \( t \) is determined by

\[
P_s(t) = q_k(t)/\sum_{j=1}^{n} q_j(t)\]

Thus, for example, the probability of choosing cutoff .3 in the second period of the task presented in Figure 1 is somewhat larger than .25.

The Parameters: A Summary

As implied by the assumptions just listed, the adapted model has 10 parameters. Six of these parameters are basic learning parameters (and are not affected by the strategy space). Erev et al. (1995) used Erev and Roth’s (1996) assessment of these parameters. The selected values—\( \sigma(1) = 3, \phi = .001, \nu = .0001, \rho(1) = 0, w^- = .02, \) and \( w^+ = .01 \)—were set to fit the matrix games data considered by Erev and Roth (1996). The remaining four parameters (\( m = 101, c_{\text{max}} = 5, \sigma_i = 1.5, \) and \( \sigma_p = .025 \)) were set by Erev et al. to address cutoff strategies (and to fit their perceptual game data).

Comparing the Model’s Predictions With Experimental Results

**Method: Computer Simulations**

As a means of comparing the model’s predictions with experimental results, computer simulations were run that were designed as a direct replication of the important (under the present model) characteristics of the experimental settings. The simulated observers “participated” in the same number of rounds as the experimental participants. One hundred simulations were run in each task. At each round of each simulation, the following steps were taken:

1. The state of the world—\( S1 \) with probability \( P(S1) \) and \( S2 \) otherwise—was randomly determined.
2. The simulated observer’s cutoff \( c_k \) was randomly determined, in accordance with Assumption 4, from the assumed set of 101 possible cutoffs.
3. The perceived signal, \( x \), was selected from the assumed normal distribution given the state of the world.
4. The simulated observer’s response was determined (\( R2 \) if and only if \( x > c_k \)).
5. Profits were calculated according to the experimental pay-off rule.
6. Propensities were updated according to the experimental pay-off.
7. The reference point for the next trial was calculated.

I turn now to a comparison of the simulation results with results that were obtained in specific studies. Each of the following sections summarizes one of the observed regularities as it is presented in the literature, along with the model’s predictions for the relevant experimental conditions. An attempt was made to organize the results in a historical order, but replications and similar results are presented together, independent of dates.

The model’s predictions are first shown without fitting parameters. That is, Erev et al.’s (1995) “original” parameters (which were set on the basis of Erev and Roth’s matrix games data and Erev et al.’s perceptual game data) are used. The main result of this analysis is that, in all cases, the model with the original parameters provides a fit to the observed data that is equal to or better than that of the (often post hoc) models presented in the original experimental articles. Additional assumptions were needed in only three cases in which experimental manipulations not modeled by Erev et al.’s quantification of the model had an effect. These effects are not predicted by alternative models and can be accounted for by fitting parameters. Finally, a sensitivity analysis is presented that shows that the model’s relative success is not a result of clever (or lucky) parameter fitting. Changes of up to 50% in the parameters’ values do not impair the model’s qualitative fit.

**Probabilistic decision process.** SDT assumes that the decision process is deterministic. That is, observers are expected to adopt a static cutoff. Lee and his associates (e.g., Lee, 1963; Lee & Janke, 1964; Lee & Zentall, 1966) examined this assumption in an external noise experimental paradigm and found robust violations. To account for these results, Lee (1963) proposed the micromatching model. Under this model, the probability of each response given a specific signal is determined by (matched to) the probability that this response will be correct (given the signal).
Representative findings (later replicated by Gilat et al., 1997; Kubovy & Healy, 1997; and Ward, 1973) were obtained in Lee and Janke's (1964) study. In this study, participants were asked to categorize stimuli (numbers, dot positions on a file card, or grayness of squares) that were samples from two normal distributions. The distance between the two distributions was $d' = 1.5$. Uniform priors were used, $P(S_1) = P(S_2) = .5$, and participants were asked to try to maximize the number of accurate decisions. Three hundred experimental trials were run.

The left side of Figure 2 shows the observed percentage of violations of the prediction of a static cutoff (referred to as static cutoff violations [SCV] $^9$) over time. Whereas stimulus type had an effect, two robust results can be observed across types. The percentage of SCV appears to be quite large (on average, above 10%), and a decrease over time is observed. As noted by Lee and Janke (1964), the results fall between the prediction of SDT (SCV = 0) and the prediction of Lee's (1963) micromatching model that implies an SCV rate of 23%.

The solid curve in Figure 2 (right) illustrates the predictions of the current (CRL) model with the original parameters. The model reproduced the experimental trend: a large percentage of SCV that decreases over time. It is easy to see that the CRL curve is closer to each of the three experimental curves than the predictions of SDT and the micromatching model. The mean squared deviation scores (between the observed and predicted percentages) for the dots, grayness, and numbers conditions were, respectively, 19, 41, and 29 for the CRL model and 127, 51, and 39 for the micromatching model.

Although the model with the original parameters did not predict the stimulus type effect, this effect can be described by the assumption that stimulus type affects one (or more) of the model's parameters. Because the stimulus type affects the initial information the participants have about the range of the possible stimuli (and the location of the two distributions), it is natural to assume that it affects the standard deviation of the distribution of the initial propensities ($\sigma_i$). When the stimuli are dots positioned on a card, the participants have complete information about the stimulus range (the cards' borders); less information is provided when the stimuli differ along a "grayness" dimension, and no range information is available when the stimuli are numbers (because all numbers are, a priori, possible). Thus, $\sigma_i$ can be expected to be relatively low in the dots condition and high in the numbers condition.

The two dashed curves on the right side of Figure 2 show that this assumption constitutes a sufficient explanation of the stimulus effect. Peaked initials ($\sigma_i = 4.5$) lead to high SCV rates (as in the numbers condition), whereas flat initials ($\sigma_i = 0.5$) lead to low SCV rates (as in the dots condition). However, other parameters can have a similar effect. Additional research is needed to compare alternative abstractions of the stimulus effect.

**Effect of experience on SCV.** Kubovy et al. (1971) found that the percentage of SCV is reduced by experience. Their experiment involved 27 experimental sessions for thousands of trials.

---

$^9$ Kubovy and Healy (1977) used the following algorithm to calculate SCV: "Suppose that there are $b$ observations in block $B$. Let $y$ take on $b - 1$ values interpolated midway between the $b$ rank ordered value of $x$ observed in block $B$. Furthermore, let $R_1(y)$ be the number of trials on which the research participant responded $R_1$ to observations that were larger than $y$, and let $R_2(y)$ be the number of trials on which the research participant responded $R_2$ to observations that were smaller than $y$. Calculate the sum $R_1(y) + R_2(y)$ for each $y$. The number of static cutoff violations is the minimum of these $b - 1$ sums" (p. 432).

$^{10}$ Lee and Zentall (1966) noted that the micromatching model provides a better fit to decisions when individuals are not informed that one-direction cutoff strategies are appropriate (as in Lee & Zentall, 1966, and Lee & Janke, 1964). Given these settings, the fits provided by the current model and the micromatching model are similar.
two conditions are presented in Table 2. Both experiments were paradigms (e.g., Gilat et al., 1997; Gopher et al., 1995). Note Thus, the present model can account for the finding of an between the two distributions (cf. Figure 1) does not change.

Figure 3 presents the estimated $d'$ in two control conditions that were run in the studies of Gilat et al. (1997) and Gopher et al. (1995). In both cases, the administrated $d'$ was 1.5. The payoff matrices and the prior probabilities that were used in the two conditions are presented in Table 2. Both experiments were run for five blocks of 100 rounds each and used a computerized version of Lee and Janke's (1964) dot position paradigm. Figure 3 (right) illustrates the predictions of the CRL model (estimated $d'$ in simulations of these conditions). The simulated participants exhibited increases in sensitivity similar to the experimental trend.

SCV in explicit and implicit cutoffs. Kubovy and Healy (1977) compared two experimental conditions in a careful examination of SCV. Condition 1 (the implicit condition) involved a replication of Kubovy et al.'s (1971) task. The prior probabilities were $P(S1) = P(S2) = .5$, and the participants earned 2 cents for an accurate guess and lost 2 cents when they erred.

Condition 2 (the explicit condition) was identical to Condition 1 with the exception that, before the presentation of the stimulus, the participants were asked to explicitly state their cutoffs. This condition was introduced to allow observation of sequential dependencies. Twelve participants were assigned to each condition. The participants in Condition 1 completed six sessions (a total of 2,400 rounds), whereas the participants in Condition 2 completed three sessions (a total of 1,200 rounds).

Figure 4 (left) shows the average number of SCV in blocks of 50 rounds in each of the experimental conditions. A significant decrease in SCV over time was observed in the two experimental conditions. Yet, as noted earlier, it seems that DMs do not converge to a static cutoff. A certain percentage of cutoff violations are observed even after vast experience.

The right-hand side of Figure 4 shows the average number of SCV performed by the simulated DMs given the original parameters and a modified set of parameters, referred to as the explicit cutoff (EC) parameters, that captures the faster learning in the explicit cutoff condition. Like the research participants, the simulated DMs reduced SCV over trials and appeared to converge to an asymptote. Before a discussion of the difference between the two conditions, it is important to note that, even with the original parameters, the CRL model outperforms the "best" probabilistic model considered by Kubovy and Healy. This model, Schoeffler's (1965) "directional generalization model," predicts an SCV rate of about 18% after experience.

Although the difference between the two experimental conditions could not be predicted on the basis of the CRL model by itself, the model can be used to understand the difference. Under the model, faster learning can be the result of a stronger recency effect or flatter initials, or both. Thus, the EC parameters are identical to the original parameters with the exception of stronger recency ($\phi = .01$) and higher initial variability ($\sigma_i = 15$).

### Table 2
The Payoff Matrices Studied by Gilat et al. (1997) and Gopher et al. (1995) and Compared in Figure 3

<table>
<thead>
<tr>
<th></th>
<th>Gilat et al. (1997): symmetric matrix, $P(S2) = .6$</th>
<th>Gopher et al. (1995): asymmetric matrix, $P(S2) = .3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State of nature</strong></td>
<td><strong>Response</strong> $S1$</td>
<td><strong>Response</strong> $S2$</td>
</tr>
<tr>
<td><strong>Response</strong></td>
<td>$S1$</td>
<td>$S2$</td>
</tr>
<tr>
<td>R1</td>
<td>.1</td>
<td>-.1</td>
</tr>
<tr>
<td>R2</td>
<td>-.1</td>
<td>.1</td>
</tr>
</tbody>
</table>

Note. Payoffs represent earnings in shekels (1 shekel = $0.33$). $S1$ = noise distribution; $S2$ = signal distribution; $R1$ = noise response; $R2$ = signal response.
be very slow in certain settings. The predicted distributions by Kubovy and Healy (1977) predict that, in the examined task, the value of the explicit condition. The demonstration that a minor quantitative difference can be responsible for the distinction between the two experimental tasks suggests, in contradic-
tion to Dorfman’s (1977) critique, that explicit and implicit cutoffs may reflect qualitatively distinct processes. Thus, the sequential trends observed in Condition 2 (and discussed subsequently) are likely to characterize behavior in the implicit condition as well.

Between-observers variability. All of the models considered by Kubovy and Healy (1977) predict that, in the examined task, observers will quickly converge to respond R2 in half of the trials (and R1 in the other half). That is, $P(R2)$ is predicted to be .5. This prediction is in line with the optimal cutoff ($c = 0$ in this case), the probability matching hypothesis, the micro-
matching model (Lee, 1963, and Schoeffler’s, 1965, variant), additive operator models (e.g., Dorfman & Biderman, 1971; Thomas, 1973), and the ideal learner model (Kubovy & Healy, 1977).

Whereas $P(R2)$ converged to .5 over participants, consistent deviations from .5 were found in individual observer behaviors. In fact, the hypothesis $P(R2) = .5$ was rejected for 20 of the 24 participants ($p < .05$).

The left side of Figure 5 shows the distribution of $P(R2)$ in the two experimental conditions. The predicted distributions under the present model with the original and the EC parameters are presented to the right of the observed distributions. Comparison of the two sets of plots reveals similar roughly single-peaked distributions. Thus, independent of the choice of parameters, the CRL reproduces the experimental trend that violates alternative models.

Learning and distance from optimal cutoff. It should be emphasized that the indication, discussed earlier, that experimental and simulated DMs do not converge to $P(R2) = .5$ does not imply that DMs do not learn in an adaptive fashion. As exemplified by Roth and Erev (1995), the learning process can be very slow in certain settings.

According to the present model, the observed deviation from .5 reflects an intermediate-term result. That is, in the long term, all simulated DMs converge to select the optimal cutoff as a modal response, and $P(R2)$ converges to .5.

Support for the predicted slow learning process can be seen in Figure 6, which illustrates the average distance of the DM’s critical point from the optimal cutoff in Kubovy and Healy’s (1977) study. The critical point is defined as the cutoff that minimizes the DM’s SCV; it is an estimation of the cutoffs used by the DMs. As can be seen in this figure, the critical points of the experimental and simulated participants move toward the optimal cutoff. Moreover, in both cases, the learning curves display the “power law of practice” (Blackburn, 1936; Cross-
man, 1959); the size of the adjustment between the first and second sessions is larger than the size of the final adjustment.

The simulated DMs were faster learners. This “extra rational-
ity” may have been a result of the assumed strategy space. Whereas the simulated DMs were confined to the range $c = -5$ to $c = 5$, the research participants could use a wider range.

Outcome effect on the direction of cutoff shifts. Recall that in the explicit condition of Kubovy and Healy (1977), cutoff shifts could be observed. Kubovy and Healy performed detailed analysis on the effect of feedback on the direction of reported cutoff shifts. As a means of evaluating error-correcting models and the more general additive operator models, the observed shifts were classified into two categories, toward and away. (Let $C_t$ be the value of the cutoff reported in trial $t$. A shift was classified as toward if $C_{t+1} < C_t$ after S1 and $C_{t+1} > C_t$ after S2. A shift was classified as away if $C_{t+1} > C_t$ after S1 and $C_{t+1} + 1 < C_t$ after S2.) Note that away shifts represent a movement in the best reply direction (i.e., an increase in the probability of the response that was correct in the previous trial) and toward shifts represent a movement in the opposite direction.

The left-hand side of Figure 7 illustrates the results of this analysis. It reveals a complex relation between the outcome of the preceding trial and the direction of the shift. Moreover, this relation changed over time.

The observed pattern is inconsistent with error correction models (that predict no shift after correct response) and with the more general additive operator models (that do not allow for toward shifts or experience effects). To account for the

![Figure 4. Kubovy and Healy's (1977) study: Frequency of static cutoff violations in blocks of 50 trials as a function of condition and time (three or six sessions of 400 trials). EC = explicit cutoff; CRL = cutoff reinforcement learning.](image-url)
results, Kubovy and Healy (1977) suggested an "ideal learner" model. This model predicts a much stronger interaction than the experimental interaction. In particular, it predicts 99% away shifts and no toward shifts after errors.

The right-hand side of Figure 7 illustrates the predictions of the CRL model. The model captured much of the complex interaction (a trend for more away shifts after errors and more toward shifts after correct responses that decreases over time). The CRL model outperformed the ideal learner model in two respects: Its quantitative predictions were closer to the observed proportions (a mean squared deviation score, between the observed and predicted percentages of less than 200 for both parameter sets relative to a mean squared deviation score of 800 obtained by the ideal learner model), and it predicted the toward shift after error that violates the ideal learner model.

Outcome and time effect on absolute cutoff shifts. Another interesting interaction between the outcome of the preceding trial and the linear trend was observed by Kubovy and Healy (1977) in an analysis of mean absolute cutoff shifts. Errors were followed by a larger cutoff shift in the first two sessions but not in the third. The left-hand side of Figure 8 shows this interaction, and the right-hand side reveals that the simulated DMs behaving in accordance with the CRL model (with both sets of parameters) exhibited a similar (although less dramatic) interaction. This interaction is inconsistent with the additive operator and ideal learner models.

No stimulus-location effect on cutoff shifts. According to the ideal learner model, shifts after a correct response should be sensitive to the location (along the c scale) of the recent stimulus. In violation of this prediction, Kubovy and Healy (1977) found no relation between the location of the stimulus and the cutoff shift. This result is consistent with the present model, which assumes that the learning process is determined by the outcome independent of the exact location of the stimuli.
The gambler's fallacy. To account for the observed toward shift after errors, Kubovy and Healy (1977) suggested that participants exhibit the "gambler's fallacy." That is, participants behave as if they expect that the state of the world will change at a probability higher than .5. Support for this suggestion comes from the observation that, after a correct response, 58% of the shifts reduce the probability of the same response. That is, under the assumption that the DMs maximized the subjective probability of a correct response (and expected payoff) in trial $t$, these shifts imply that the positive feedback ("the response was correct") decreased the subjective probability that this response would be correct again at trial $t + 1$.

The CRL model predicts this trend. The proportion of shifts that reduced the probability of repetition after a correct response in the simulation was 60%. Thus, the simulated DMs behaved as if they had biased expectations. Of course, these simulated individuals had no "expectations" at all. Careful examination of the reason for the apparent gambler's fallacy reveals that it was a result of a regression effect. Under the present probabilistic model, the probability of a shift toward the central cutoff ($c = 0$) is typically larger than .5.

Conservative cutoffs. Another phenomenon that can be explained by a regression effect is conservatism (see Erev, Wallsten, & Budescu, 1994). In the present context, conservatism can also be a result of a slow learning process.

Conservatism is defined relative to an optimal response rule. The optimal (likelihood ratio) response rule in signal detection tasks can be expressed as "response $R_2$ when the likelihood ratio of the stimulus—$P(x|S_2)/P(x|S_1)$—exceeds the optimal cutoff." Assuming risk neutrality, the optimal likelihood ratio cutoff, typically referred to as $\beta^*$, is determined by the payoff matrix and the prior probabilities. When the payoffs and the priors are symmetric (as in Kubovy & Healy, 1977), $\beta^* = 1$. Namely, in Figure 1a, the cutoff is located at $c = 0$. Asymmetry moves $\beta^*$ away from the center of the two distributions.

Experimental investigation of asymmetrical signal detection tasks reveals that human DMs tend to select a cutoff ($\beta$) between $\beta^*$ and 1. This phenomenon, referred to as conservative cutoff placement, has been replicated in many studies (at least since Green & Swets, 1966/1988).

Conservative cutoff placement and related phenomena were demonstrated and studied by focusing on Experiment 1 of Healy and Kubovy (1981). This experiment used Kubovy et al.'s (1971) paradigm to compare competing explanations for the conservative cutoff placement phenomenon. Participants completed 60 trials under each of the three payoff matrices by three prior probability conditions. Each experimental point increased the participants' payoff by 1 cent.

The observed experimental and simulation results in the nine conditions are displayed in Figure 9. The horizontal axis shows the optimal cutoffs, and the vertical axis shows the observed cutoffs (logarithmic scales). As can be seen, both the experimental and simulated DMs exhibited conservative cutoff placement.

Probability matching. Among the explanations compared by Healy and Kubovy (1981), the additive probability matching rule was best supported by the results. According to this account, observers behave as if trying to match the proportion of $R_2$ responses to $P(S_2)$ plus a constant. The value of the constant
is determined by the payoff matrix. When the payoff matrix is symmetrical (as in Table 2, left), the constant is zero, and \( P(R2) \) should equal \( P(S2) \).

Yet, clear violations of this rule were observed in Healy and Kubovy's (1981) Experiments 2 and 3 (not summarized here). Healy and Kubovy (1981) concluded that "although the additive probability matching rule fares better than any other explanation for conservative cutoff placement that has been proposed, it is no more than a first approximation of cutoff location in probabilistic categorization tasks" (p. 353).

According to the present approach, probability matching is an intermediate-term result rather than a general principle that guides behavior. The left side of Figure 10 shows the proportions of \( R2 \) responses in two control conditions that were run (for five blocks of 100 trials each) in the Gilat et al. (1997) study. Both conditions used Table 2's symmetric matrix (payoffs of .1 shekel for a correct response and -.1 shekel for an error) with \( P(S2) = .6 \). The experiments involved a computerized version of Lee and Janke's (1964) dot position paradigm with administrated \( d' \) values of 1 and 1.5. Figure 10 shows that both the experimental and the virtual DMs appeared to engage in probability matching after 100 rounds but slowly moved toward the optimal higher proportion of \( R2 \) choices.

Relative effect of payoffs and prior probabilities. The optimal (risk-neutral) cutoff is equally sensitive to the payoff ratio and the prior probabilities likelihood ratio. Specifically, \( \log(\beta^*) = \log(\text{payoff ratio}) + \log(\text{prior ratio}) \).

Healy and Kubovy (1981) found that their research participants tended to be more sensitive to the priors than to the payoffs ratio. For example, their Experiment 1 data can be summarized by the following regression equation: \( \log(\text{observed } \beta) = .01 + .19 \log(\text{payoff ratio}) + .3 \log(\text{prior ratio}) \). The virtual DMs showed a similar "sensitivity" pattern. The regression equation that summarizes the simulation results (with the original parameters) is \( \log(\text{simulation } \beta) = .004 + .16 \log(\text{payoff ratio}) + .2 \log(\text{prior ratio}) \).

To understand the intuition behind this result, consider a specific example (one of the conditions studied by Healy and Kubovy, 1981) in which a hit = 3 (payoff for accurate \( R2 \)), a correct rejection = 1 (accurate \( R1 \)), a miss = false alarm 0, and \( P(S1) = .75 \). Whereas the optimal payoff is \( \log(\beta^*) = \log(3/1) + \log(.25/.75) = 0 \), the average (human and virtual) DM tends to set positive cutoffs that imply more than 50% \( R1 \) choices. Examination of the behavior of the virtual DMs reveals that their behavior was driven by the large effect of the very first reinforcements. Positive cutoffs are reinforced with higher probabilities (because \( S1 \) is the more likely state); thus, the first reinforcement is more likely to be given to one of these cutoffs. And since the learning process becomes very slow after the first few trials, these initial reinforcements have long-term effects.

This observation implies that, under the current model, the relative overweighting of the prior odds is not a general phenomenon. An opposite effect can be predicted given distinct payoff matrices. For example, simulation results reveal that a subtraction of 3 from all payoffs in the current example is expected to lead to underweighting of the priors.

Note that overweighting priors when all payoffs are positive (the original example) "looks like" risk aversion, a bias toward the alternative with smaller payoff variance. And underweighting priors in the loss domain (the modified example) looks like risk seeking. Thus, the current finding reinforces March's (1996) assertion that the reflection effect (Kahneman & Tversky, 1979) can be a result of a reinforcement learning process.

The current interpretation of Healy and Kubovy's (1981) results can also help explain the apparent inconsistency between these results and Green and Swets's (1966/1988) observation of a stronger payoff effect. Green and Swets used payoff matrices with both positive and negative outcomes. Yet, other explanations for these inconsistencies are, of course, possible. One reasonable explanation involves the sample size: Green and Swets's study had only 1 participant.
**Initial maladaptive learning.** A particularly strong violation of SDT has been observed in certain settings in which experience appears to impair the quality (in terms of expected value) of the chosen cutoffs. In these settings, DMs do not update their cutoffs in the direction predicted by the theory; they appear to move their cutoffs away from the optimal point.

For example, in the control condition of the Erev et al. (1995) study, participants made 200 decisions given Table 2's asymmetric payoff matrix. The participants were asked to distinguish between two letters (P [S1] and F [S2]) that were presented for 60 ms (the average estimated \(d'\) was above 3). The prior probability, \(P(S2)\), was .3. Whereas the optimal \(\beta\) is .233, \(\log(\beta^*) = -1.46\), the observed \(\beta\)s were above 1, \(\log(\text{observed } \beta) > 0\).

A similar trend was observed in an external stimulus replication of that study (Gopher et al., 1995; also summarized in Erev & Gopher, in press) in which the stimuli were dot positions. Gopher et al.'s study was longer (500 rounds) and included a manipulation of the administrated \(d'\) (1.5 and 2.5). The results (Figure 11, left) reveal learning in the "wrong" direction in the first 100 periods given a high administrated \(d'\). More experience moves participants toward the optimal cutoff. Similar trends were observed in the simulation of the CRL model.

**Effect of \(d'\) on the learning process.** Another interaction between the effects of administrated \(d'\) and experience was observed by Healy and Kubovy (1977). Experiment 2 of that study compared the effect of prior probabilities on cutoff locations in a recognition memory (discussed subsequently) and two numerical decision tasks (Kubovy et al.'s 1971 paradigm). Two \(d'\) levels were studied: \(d' = 0.5\) and \(d' = 1.5\). Two prior probability levels—\(P(S2) = .5\) and \(P(S2) = .25\)—were compared given symmetric payoff matrices; the optimal cutoffs are \(\beta^* = 1\) in the symmetric case and \(\beta^* = 3\) when \(P(S2) = .25\). Participants completed four blocks of 40 rounds in each condition.

The general trends predicted by SDT were observed in the two conditions: With uniform priors, participants learned to converge to \(\beta = 1\), whereas asymmetrical priors led to higher cutoffs. Yet, a clear difference between the two experimental conditions was observed. As can be seen in Figure 12, given asymmetric priors, the distance between the optimal and the observed cutoffs was affected by the experimental condition. The simulated participants exhibited the same trend.

**Difference between internal and external paradigms.** Healy and Kubovy (1977) compared the two external conditions just discussed with a recognition memory task in which participants were asked to judge whether a given number had been shown previously. Under SDT, this condition involves internal error distributions.

A slower learning process was observed in the recognition task. This finding cannot be a result of the sensitivity effect (discussed earlier), because the estimated \(d'\) in the recognition task was between the estimated \(d'\) values in the two external tasks. Under the present model, this finding can be described by the assumption of a smaller initial propensity variability \((\sigma_t)\) in the internal paradigm. That is, participants come to the laboratory with a strong tendency for an unbiased cutoff in memory tasks. Because DMs are likely to have less experience with numerical tasks, their initial propensities are flatter. As noted earlier, flatter initials speed the learning process.

**Payoff order (or variability) effect.** Busemeyer and Myung (1992) ran an experiment (Experiment 2 in their article) to compare hill climbing and error correcting models in an explicit cutoffs paradigm.\(^{12}\) The experimental task was similar to Kubovy and Healy's (1977) Condition 2 task, with two important exceptions. First, the research participants were restricted to 101 cutoffs uniformly distributed in the range \(c = -3\) to 3. Second, the number of stimuli categorized by each stated cutoff decision was manipulated. Three sample size conditions were compared: 1, 3, and 15. The condition involving a sample size of 1 was similar to Condition 2 of Kubovy and Healy (1977); each stated cutoff was used to categorize a single stimulus. In the other two conditions, 3 and 15 stimuli were categorized by each stated cutoff.

\(^{12}\) The hill climbing model is only a (noncentral) part of the theoretical approach suggested by Busemeyer and Myung (1992). As noted earlier, their general approach is similar to the approach taken here.
The two payoff matrices shown in Table 3 were compared given uniform prior probabilities, \( P(S_1) = P(S_2) = 0.5 \). Participants started the experiment with 10,000 points (convertible to money) and could earn or lose points based on their performance. Whereas SDT and error correcting models predict identical cutoffs in the two matrices, the hill climbing model predicts an interaction between the payoff matrix and the sample size condition. The predicted "payoff matrix effect" is a result of the fact that the assumed learning process (given the parameters used by Busemeyer & Myung, 1992) is sensitive to outcome rank ordering (rather than to cardinal values). As can be seen in Table 3, the two matrices have a distinct outcome ordering.

Figure 13 shows the observed results, the predictions of the hill climbing model, and the predictions of the CRL model (with the original parameters). In line with the experimental manipulation, the cutoffs used by the virtual observers were limited to the range \( c = -3 \) to 3. Following Busemeyer and Myung (1992), the reported statistics are proportions of cutoffs above \( c = 0 \) (the optimal cutoff is \( c = 0.925 \) \( \beta^* = 4 \)) in three asymmetric blocks (1-10, 11-50, and 51-150). Busemeyer and Myung noted that the correlation between the 18 observed proportions and the predictions of the hill climbing model was 0.84. As it happened, 0.84 was also the correlation between the CRL model predictions and the observed proportions.

Note that the present model is not sensitive to outcome rank ordering. The predicted slower learning process in Matrix 2 with a sample size of 1 is a result of the higher variance of the payoffs in this matrix. Thus, the present results demonstrate that the difference between the two matrices observed in the experiment may not reflect an ordinal learning process. Future research should compare the ordinal explanation (implied by Busemeyer and Myung’s model) with the present variability explanation.

The trend in one of the six experimental curves (Matrix 2, sample size of 1) is not captured by the present model. Participants in this condition moved toward the suboptimal range during the 150 experimental trials. The simulated participants moved toward the suboptimal range during the first block (in which the proportion of correct range decisions was below 0.5) but learned in an adaptive fashion in the last two blocks. This deviation may reflect a learning modeling problem, but it can also be the result of an inappropriate model of the incentive structure. In this condition, participants could have been in a state of heavy losses. Assuming that the research participants could not be asked to pay their losses, these losses would be expected to affect the incentive structure (losses above a certain point have no meaning).

**Effect of opportunity to revise decisions.** Figure 13 demonstrates that, in Busemeyer and Myung’s (1992) study, giving the participants the opportunity to revise their decisions could have impaired the learning process. Compare, for example, the result of the first 50 trials with a sample size of 3 and the result of all 150 trials with a sample size of 1. In both cases, the

---

**Table 3**

*The Payoff Matrices Compared by Busemeyer and Myung (1992) and Analyzed in Figures 12 and 13*

<table>
<thead>
<tr>
<th></th>
<th>Matrix 1</th>
<th>Matrix 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>S1</td>
<td>S2</td>
</tr>
<tr>
<td>R1</td>
<td>50</td>
<td>-400</td>
</tr>
<tr>
<td>R2</td>
<td>-100</td>
<td>200</td>
</tr>
</tbody>
</table>

*Note.* S1 = noise distribution; S2 = signal distribution; R1 = noise response; R2 = signal response.
participants categorized 150 stimuli and received the same amount of information. The only difference between these two cases involved the opportunity to revise decisions. Participants could revise their cutoff in every trial involving a sample size of 1 but only every three stimuli with a sample size of 3. The results revealed a higher percentage of decisions in the "correct" direction (cutoffs above c = 0) when cutoffs could not be revised after each observation.

This counterintuitive phenomenon is a violation of error correcting models but, as can be seen in Figure 13, is predicted by the hill climbing model and by the present model. Both models predict slower learning with a sample size of 1.

Backward hill climbing. According to the hill climbing model, the cutoff shift in trial t (the difference between the cutoff in trial t and the cutoff in trial t - 1) is determined by the success of the shift in trial t - 1. If the previous shift was successful (outcome in trial t - 1 larger than outcome in trial t - 2), another shift in the same direction is expected. An unsuccessful shift in trial t - 1 is expected to lead to a shift in the opposite direction in trial t.

Figure 14 shows the proportions of observed repeated shifts as a function of the success of previous shifts in Experiment 2 of Busemeyer and Myung (1992; the data are weighted averages of Busemeyer & Myung's Table 4 data). The data show no support for the hill climbing prediction after successful shifts. In violation of the hill climbing model, fewer than 50% of these shifts (only 48.7%) were in the predicted direction. The present model reproduces the experimental trends, including "backward hill climbing" after successful shifts.

Summary and Sensitivity Analysis

Table 4 presents a summary of the 19 phenomena just discussed. The first column summarizes each of the observed trends (note that some of the 19 phenomena are summarized by more than one trend). The second column indicates whether the CRL model with the original parameters (and with the EC parameters for Kubovy and Healy's, 1977, explicit cutoff condition) can account for the observed qualitative trends. This column provides positive answers in all but three cases in which additional assumptions (parameters changes) were needed.

Before discussing the model's success, it is important to consider the three failures and their implications. Note that all three cases involve response mode and display mode (stimulus type) effects. The current model can account for these effects by "mode-specific" parameters but cannot predict them. Additional research is needed to develop a theoretical framework for predicting such effects. Yet, two observations suggest that the value of the current model is not conditional on the resolution of the response and display mode effects. First, even without the additional assumptions, the model outperforms alternative explanations of the relevant phenomena. Second, the additional effects do not appear to interact with the other observed effects reproduced by the model independent of the parameter changes; for example, both the original parameters and the EC parameters reproduce the main trends observed by Kubovy and Healy (1977).

Qualitative comparisons between the CRL model and alternative models can be obtained by comparing columns 1 and 2 in Table 4. Column 1 indicates that some of the qualitative trends are violations of specific models. In fact, all of the alternative models considered here are violated by at least one observed trend. Column 2 shows that these trends are in line with the current model.

Quantitative model comparisons are presented in the third column of Table 4. This column summarizes a comparison of the model's quantitative predictions and the predictions of alternative quantitative models (that are not violated by the data's qualitative trends). Alternative quantitative models are available in six cases. In four of them, the current quantitative predictions are closer to the data than the alternative models. In the other two cases, the fit provided by the CRL model is comparable to the fit provided by the hill climbing model.

The fourth column of Table 4 presents the results of a sensitivity analysis evaluating the robustness of the model's predictions to variations in the values of each of the models' nine nonzero original parameters: m, C_max, s(1), phi, u, w^-, w^+, o_j, and sigma_j. Two sets of an additional 100 simulations were run to evaluate the effect of each parameter (when all other parameters are fixed). In one set, the parameter's original value was decreased by 50%; in the other set, it was increased by 50%. Thus, 18 additional sets of simulations (9 parameters \times 2 variations) were run in each of the experimental tasks. Column 4 indicates which of the 50% parameter variations impair the model's qualitative predictions in each case. A variation is said to impair prediction if it leads to a prediction that is inconsistent with the trend described in the first column. Column 4 shows that the main qualitative predictions are robust to almost all of the examined parameter variations. Note that both exceptions (low C_max and low m) involve a reduction of the assumed strategy space. This finding implies that the model's relative success is not a result of the choice of parameters. Rather, it seems that the assumption of reinforcement learning among cutoff strategies has robust implications, and these robust implications, rather than the specific parameters, are supported by the current research.

It is important to note, however, that the fact that the qualitative predictions are relatively insensitive to the specific parameter modifications appears to be "reasonable" (e.g., unknown boundaries increase initial variability and explicit statement of the cutoff increases recency) suggests that general principles can be found.

---

13 The observation that the necessary parameter modifications appear to be "reasonable" (e.g., unknown boundaries increase initial variability and explicit statement of the cutoff increases recency) suggests that general principles can be found.
Table 4
Summary of the 19 Phenomena and Observed Trends

<table>
<thead>
<tr>
<th>Phenomenon</th>
<th>Qualitative fit with original parameter (+) or with additional assumptions (AA)</th>
<th>Comparison with alternative quantitative models</th>
<th>Sensitivity parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCV (Lee &amp; Janke, 1964)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between 0 and micromatching</td>
<td>+</td>
<td>&gt;Lee (1963)</td>
<td>None</td>
</tr>
<tr>
<td>Stimulus effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience effect on SCV (Kubovy et al., 1971)</td>
<td>+</td>
<td>&gt;Lee (1963)</td>
<td>None</td>
</tr>
<tr>
<td>Estimated d' (Gilat et al., 1997; Gopher et al., 1995)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term matrix effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response mode effect on SCV (Kubovy &amp; Healy, 1977)</td>
<td>AA (e.g., EC parameter)</td>
<td>&gt;Schoeffler (1965)</td>
<td></td>
</tr>
<tr>
<td>Between-observers variability (violation of all models considered by Kubovy &amp; Healy, 1977)</td>
<td>+ (EC +)</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Learning and distance from optimal cutoff (Kubovy &amp; Healy, 1977)</td>
<td>+ (EC +)</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Outcome effect on direction shifts (Kubovy &amp; Healy, 1977)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three-way interaction (violation of EC models)</td>
<td>+ (EC +)</td>
<td>Ideal learner</td>
<td>None</td>
</tr>
<tr>
<td>Away shift after correct response (violation of ideal learner)</td>
<td>+ (EC +)</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Outcome and time effect on absolute cutoff shifts (Kubovy &amp; Healy, 1977)</td>
<td>+ (EC +)</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>No stimulus-location effect on cutoff shifts (violation of ideal learner; Kubovy &amp; Healy, 1977)</td>
<td>+ (EC +)</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>The gambler’s fallacy (Kubovy &amp; Healy, 1977)</td>
<td>+ (EC +)</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Conservative cutoffs (Healy &amp; Kubovy, 1981)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability matching (and its violation)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term trend (Healy &amp; Kubovy, 1981)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time effect (Gilat et al., 1997)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative effect of payoffs and prior probabilities (Healy &amp; Kubovy, 1981)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maladaptive learning (Erev et al., 1995; Gopher et al., 1995)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term negative effect given high d'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long term</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of d' on the learning process (Healy &amp; Kubovy, 1977)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning speed difference between internal and external paradigms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Healy &amp; Kubovy, 1977)</td>
<td>AA (e.g., initial parameter)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payoff order (or variability) effect (Busemeyer &amp; Myung, 1992)</td>
<td>+ (5 of 6)</td>
<td>= Hill climbing</td>
<td>None</td>
</tr>
<tr>
<td>Effect of opportunity to revise decisions (Busemeyer &amp; Myung, 1992)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backward hill climbing (violation of hill climbing; Busemeyer &amp; Myung, 1992)</td>
<td>+</td>
<td>= Hill climbing</td>
<td>None</td>
</tr>
</tbody>
</table>

Note. SCV = static cutoff variations; EC = explicit cutoff.  
* 50% change in sensitivity parameter value impairs the model’s reproduction of the observed trends.

ter values does not imply that these parameters can be ignored. Simple thought experiments reveal that positive values of these parameters are needed to ensure adaptive learning. For example, both m and Cmax have to be positive to allow learning, and when the reference point does not move (w- = w+ = 0), no learning can occur in the loss domain.

Discussion

The results presented here demonstrate that a minimal modification of SDT, replacement of the ideal observer rational response rule with a CRL rule, is sufficient to account for the robust violations of this theory. The modified (CRL) signal detection model appears to capture a general principle that is not captured by alternative descriptive variants of SDT. It accounts for the probabilistic nature of binary categorization better than previous probabilistic models (Lee, 1963; Schoeffler, 1965) and accounts for sequential dependencies better than models that assume cutoff direction learning (e.g., Busemeyer & Myung, 1992; Dorfman & Biderman, 1971; Kubovy & Healy, 1977; Thomas, 1973). In none of the 19 behavioral regularities considered here were the model’s predictions outperformed by the (often post hoc) explanations proposed in previous articles. This improved approximation was achieved without parameter fitting and appears to be relatively robust to choice of parameters.

Robustness of the Processes Underlying Categorization Decisions

The experimental regularities summarized by the present model have been observed in a wide range of experimental paradigms. The distinct paradigms differ along nine main dimensions: error source (external, internal, or combined), stimulus type (e.g., dot positions, numbers, grayness, letters, and memories), response mode (binary decisions or explicit cutoffs), payoff information (known or unknown payoff matrix), prior
probability information (known or unknown priors), informative value of the feedback (only payoffs or feedback that includes or implies the correct state of nature), incentives (hypothetical or cash profits), length (from 60 to more than 3,200 trials), and strategic complexity (decision under uncertainty or complex games). The observation that one simple model can reproduce the observed behavior in all paradigms implies that the model approximates an extraordinarily robust psychological process. This conclusion is consistent with the view expressed by Kubovy and Healy (1977; see also Healy & Kubovy, 1977). On the basis of careful experimental examinations of a subset of the nine dimensions considered here, they suggested that a robust categorization process exists. As noted in the introduction, this conclusion is also expected from ecological considerations given the high frequency and potential survival value of decisions of this type.

It is important to emphasize that the good approximation provided by the model does not imply that the model (with the original parameters or with any other set of parameters) is exactly accurate; nor does it imply that factors not explicitly modeled here (such as prior information concerning the payoff rule and prior probabilities and the informative value of the feedback) have no effect. Rather, the results suggest that the model captures robust aspects of binary categorization decisions, aspects that explain a large portion of the variability within a wide set of situations.

Specific Practical and Methodological Implications

SDT is one of the best examples of a psychological theory with nontrivial practical implications. Among the subfields in which it is used are data analysis (Macmillan & Creelman, 1991), human factor engineering (Davis & Parasuraman, 1982), medical decision making (e.g., Swenssen et al., 1977), group decision making (e.g., Sorkin & Dai, 1994), statement categorization (Wallsten & Gonzalez-Vallejo, 1994), and probability assessments (e.g., Budescu, Wallsten, & Au, 1997; Ferrel & McGoey, 1980). Yet, as noted in the introduction, the apparent success of applications of this theory appears to be inconsistent with the observation that the theory is violated in experimental tests. The current research provides a resolution to this contradiction. Under the CRL signal detection model, applications of SDT can be successful (in many settings, the predictions of the CRL model are close to the optimal cutoff), and certain violations of SDT are predictable.

This resolution has two types of practical implications. First, it can be used to evaluate the conditions under which the applications of traditional SDT are likely to be robust. For example, the current model predicts that the comparison of estimated $d'$ across tasks is likely to reflect true relative perceptual abilities when the payoff matrices are similar. Under the current model, estimated $d'$ is a function of the "true" $d'$ and the incentive structure. Thus, when the incentives are fixed, the correlation between estimated $d'$ and true sensitivity should be high.

A second, more interesting implication involves an attempt to use the current results to improve available applications. For example, the static cutoff assumption (an assumption implicit in the calculation of SDT statistics) can, in principle, be replaced by an approximation of SCV that can be provided by the present model.

Another set of new implications involves complex situations (e.g., $n$-person signal detection games) for which the ideal observer's prescription cannot always be obtained. As long as the incentive structure is clear, the present model can be used to describe behavior (even when optimal behavior is not known).

Value of the Cognitive Game Theoretic Approach

The current investigation used a cognitive game theoretic decomposition of SDT to facilitate identification and replacement of the weak elements of this theory. The decomposition was first used to identify one submodel of SDT (the ideal observer decision rule) that is likely to be violated (by generalizing from judgment and decision-making research). I then generalized from matrix game research and suggested a reinforcement learning replacement to this submodel. Finally, I built on Erev et al.'s (1995) investigation of perceptual games to set initial parameters for the new submodel.

The relative success of the revised model provides another demonstration of the value of the cognitive game theoretic approach. As suggested by previous research (Erev & Roth, in press; Gilat et al., 1997; Roth & Erev, 1995), the cognitive game theoretic decomposition appears to facilitate generalization across tasks. The accumulated evidence illustrates that the distinctions among the incentive structure, the cognitive strategy space, and the choice-learning rule help clarify three robust principles.

1. In a wide set of repeated decision tasks (including matrix games and binary categorization decisions), DMs are sensitive to objective incentives and information structure (money and likelihood ratios). Thus, SDT abstraction of the information and payoff matrix is useful.

2. In a wide set of situations (including market entry decisions [studied in Rapoport et al., in press] and binary categorization decisions), the cognitive strategies considered by DMs can be approximated by a set of cutoff strategies (in line with SDT).

3. In an even wider set of repeated decision tasks (including matrix, extensive form, market entry, and team games, along with binary categorization decisions), DMs violate the predictions of rational decision theory but nevertheless slowly adjust their selected cognitive strategies in response to the objective incentive structure. This adjustment process can be approximated by Roth and Erev's (1995) linear quantification of the law of effect (and the process may also be approximated by other slow adaptive learning processes).

Note that each of the three principles just stated is almost trivial. The suggestions that people adjust their strategies in response to the incentive structure (Principles 1 and 3) and that cutoff strategies are often used (Principle 2) do not appear controversial (or interesting). Yet, when the three principles are quantified, they lead to nontrivial predictions. And, in the current setting, these predictions are consistent with 19 robust behavioral regularities.

Implications for Direction Learning Models

The apparent advantage of the CRL model over direction learning models (the linear operator, error correcting, and hill
Implications for Judgment and Decision-Making Research

Although the current research appears to be quite remote from contemporary judgment and decision-making research (which focuses on one-period decision tasks), some important connections exist. Note, first, that cognitive game theoretic analysis of choice behavior is possible only when the cognitive strategy space can be approximated. Thus, this analysis must be based on basic judgment and decision-making research to approximate the strategies (heuristics) that DMs tend to use. Whereas the current research was built on rather simple cognitive strategies (cutoff strategies), analysis of more complex choice situations will have to rely on the abstraction of more complex heuristics.

At least in the context of repeated decisions, the current approach can complement basic judgment and decision-making research. One of the main difficulties in applying judgment and decision-making research arises from the observation that people tend to follow more than one heuristic in relatively similar situations. To address this difficulty, Payne, Bettman, and Johnson (1993) proposed the adaptive DM framework. Under this framework, DMs tend to follow adaptive rules that maximize accuracy and minimize effort. The cognitive game theoretic approach can be thought of as an extension of this idea. It can be used to model the process by which DMs become adaptive and can address situations in which other incentives (in addition to accuracy and cognitive effort) may be important.

Finally, the current results shed light on the apparent contradiction between the “heuristic and biases” (e.g., Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 1973, 1996; Tversky & Kahneman, 1974) and ecological (e.g., Gigerenzer & Hoffrage, 1995) approaches to the study of human judgment and decision making. Whereas the heuristic and biases research demonstrates that human judgment can be approximated by a limited set of (typically) adaptive cognitive strategies (that can lead to biases), the ecological research demonstrates that, in certain “ecological” settings, people behave as if they are “frequentist statisticians.” Clearly, both types of observations can be consistent with the current view. As the present results demonstrate, the fact that people learn among adaptive strategies does not imply that they will not be biased. Yet, given a certain incentive structure, bias-free behavior is possible.

Conclusion

The present result, the demonstration that a simple cognitive game theoretic variant of SDT can account for the main behavioral regularities observed in binary categorization decisions, has two important implications. First, it suggests that the simple principles underlying this model are likely to characterize categorization decisions in a wide set of situations. According to these principles, behavior can be predicted by the assumption that DMs follow cutoff strategies and slowly adjust their cutoffs in response to the incentive structure. Moreover, the observed behavior is relatively robust; one quantification of these principles provides a good approximation of behavior in all of the experimental tasks studied here.

A second implication is related to the potential of the cognitive game theoretic approach. The success of the current example of a game theoretic model and the success of similar models in different settings (in particular, Erev & Roth’s, in press, account of behavior in all experimental matrix games with mixed strategy equilibrium) demonstrate the potential of the general approach. Future research should examine whether other classes of cognitive or decision phenomena can be predicted by a cognitive game theoretic analysis.

References


Thomas, E. A. C. (1973). On a class of additive learning models: Error


Received December 5, 1995
Revision received August 11, 1997
Accepted August 11, 1997