ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 74, No. 2, May, pp. 118–144, 1998
ARTICLE NO. OB982772

Accidents and Decision Making under Uncertainty: A Comparison of Four Models

Rachel Barkan, Dov Zohar, and Ido Erev

Technion—Israel Institute of Technology, Haifa, 32000, Israel

Heinrich's (1931) classical study implies that most industrial accidents can be characterized as a probabilistic result of human error. The present research quantifies Heinrich's observation and compares four descriptive models of decision making in the abstracted setting. The suggested quantification utilizes signal detection theory (Green & Swets, 1966). It shows that Heinrich's observation can be described as a probabilistic signal detection task. In a controlled experiment, 90 decision makers participated in 600 trials of six safety games. Each safety game was a numerical example of the probabilistic SDT abstraction of Heinrich's proposition. Three games were designed under a frame of gain to represent perception of safe choice as costless, while the other three were designed under a frame of loss to represent perception of safe choice as costly. Probabilistic penalty for Miss was given at three different levels (1, .5, .1). The results showed that decisions tended initially to be risky and that experience led to safer behavior. As the probability of being penalized was lowered decisions became riskier and the learning process was impaired. The results support the cutoff reinforcement learning model suggested by Erev *et al.* (1995). The hill-climbing learning model (Busemeyer & Myung, 1992) was partially supported. Theoretical and practical implications are discussed. © 1998 Academic Press

In his classical study, Heinrich (1931) noted that most industrial accidents occur as a result of an unfortunate conjunction of a human error and a chance event. In Heinrich's sample of 5000 documented industrial accidents the likelihood ratio of an accident, given a human error (i.e., failure to detect an unsafe

This research was supported by the Committee for Research and Prevention in Occupational Safety and Health, Israel Ministry of Labor and Social Affairs. It was presented at the 29th Annual Meeting of the Society for Mathematical Psychology, Chapel Hill, North Carolina, August 1996.

Address correspondence and reprint requests to be addressed Dov Zohar, Faculty of Industrial Engineering and Management, Technion—IIT, Haifa 32000, Israel.

0749-5978/98 $25.00

state) was about 1:300. More recent studies (cf. Starbuch & Milliken, 1988; Petersen, 1988; Reason, 1995) reinforce Heinrich's argument. It seems that accidents are affected by a combination of environment (safe or unsafe) and the choice between a *safe* response (preventing an accident) and a *risky* response (possibly resulting in an accident). For example, consider an operator in a chemical plant, who controls a heating process of certain chemicals. If the chemical container appears to be overfilled, the operator can take a safe (yet costly) action and empty it or a risky action and start the heating process. Starting the heating process when the container is overfilled may result in an explosion. Yet, Heinrich's observation states that in most cases a failure to detect an overfilled container does not lead to an accident (this example represents an actual problem that was studied recently by Gopher & Barzilai, 1993).

Much of the safety research that followed Heinrich's observation focused on the organizational rather than the individual level of decision making and the acquisition of risky habits. To describe the organizational level, Heinrich (1931) proposed the Domino model; i.e., an accident is an end result of a chain of decision fallacies. The Domino model offered useful insights regarding the conflict between safety and productivity, the relation of accidents to general policies and day-to-day management fallacies, and the interdependency of individual–organization–technology (see Reason, 1995; Petersen, 1988; McKenna, 1988). Neither recent versions of the Domino model (Reason, 1995; Petersen, 1988) nor motivational approaches to risk-seeking and risk-avoidance (Wilde, 1988; Summala, 1988) can be used to derive quantitative predictions. Faced with these limitations Moray (1990) made an explicit "plea for the use of quantitative models. . . of the human cognitive system" (p. 1212).

The present paper builds upon knowledge that has been accumulated in behavioral decision making research, in an attempt to address Moray's suggestion. It proceeds as follows: Section 1 presents a quantification of Heinrich's basic observation. It shows that the choice between safe and risky actions can be modeled as a probabilistic signal detection task (Green & Swets, 1966). We present six numerical examples (i.e., payoff conditions) of this model. Section 2 reviews four behavioral models providing a quantitative prediction of behavior for the numerical examples (payoff conditions) under investigation. Two of the four models we consider provide static predictions for choice behavior, reflected in a single cutoff. The first is the ideal observer model implied by traditional signal detection theory (SDT) under the assumption of expected utility maximization. The second model is an adaptation of prospect theory (Kahneman & Tversky, 1979) reflecting systematic bias from the maximization rule, which is referred to here as the PT observer. In most natural settings, adequacy of decisions is learned from experience. The two other models we consider provide dynamic predictions, reflected in adaptive changes of cutoff selection. These include Busemeyer and Myung's (1992) hill-climbing (HC) learning model and Erev, Gopher, Itkin, and Greenshpan's (1995) cutoff reinforcement learning (CRL) model. Whereas all four basic models have been supported by previous research, their application to the current context yields conflicting predictions. An experiment designed to compare the four models is presented in Section 3, and theoretical and practical implications are discussed in Section 4.

1. A QUANTIFICATION OF HEINRICH'S BASIC SAFETY PROBLEM

Analysis of 5000 industrial accidents led Heinrich (1931) to propose three elements that define the basic safety problem. The first element is the initial environment which consists of two general states—a safe or an unsafe state. The second element is the decision space that involves choice between safe and risky acts. The individual is assumed to diagnose the state of the environment as safe or unsafe and to choose a safe or a risky act. Human error (failure to detect an unsafe state) can lead to an accident. Heinrich's analysis indicated, however, that most human errors do not result in any harm or loss. Given human error, an accident occurred once in 300 pairings of unsafe conditions and risky acts. Hence, the third element proposed by Heinrich (1931) was the probabilistic nature of an accident given human error.

We suggest that the basic safety problem as Heinrich put it can be modeled as a signal detection task with the unique characteristic of having a probabilistic penalty. Heinrich's first two elements (i.e., two states of the environment and two possible choices) are summarized by the traditional SDT's payoff matrix. In safety problems the two states (Signal and Noise in SDT) are referred to as unsafe and safe conditions, and the two responses ("Signal" and "Noise" in SDT) are referred to as safe and risky acts. The third proposition made by Heinrich is not modeled by traditional SDT, but can be added to the model. This element implies that the Miss outcome (the conjunction of an unsafe state and a risky act) is a gamble. The frequent outcome of that gamble is a near accident (i.e., Miss error did not result with any damage), the less frequent outcome of that gamble is an accident. In Heinrich's study this gamble had an accident as a defined outcome in 1 out of 300 cases. Table 1 presents a variant of the SDT payoff matrix representing Heinrich's principles. Note that the traditional Miss cell is replaced with a three-parameters gamble: *P,* probability of accident given Miss; Near, the outcome of a "near accident"; Accident, conjunction of Miss and "bad luck" leading to accident.

SDT implies that the information available to the decision maker in a binary decision task with two possible states of nature (as in Table 1) can be represented by a unidimensional signal that is sampled from one of two distributions

TABLE 1

Notations for Probabilistic Penalty Signal Detection

	State of environment	
Decision	Unsafe (Signal)	Safe (Noise)
Safe act ("Signal")	Hit	False alarm
Risky act ("Noise")	$\begin{Bmatrix} \text{(Accident, } P\text{)} \\ \text{(Near accident, } 1 - P\text{)} \end{Bmatrix}$	Correct rejection

(states of nature). In fact, even when the environmental stimuli are multidimensional they are assumed to be reducible to a single dimension that represents the relative likelihood of the two states of nature. Figure 1 presents an example in which the two distributions are assumed to be normal with equal variance. The distance between the means of the two distributions (in standard deviation units), referred to as d', is a measure of the observer's sensitivity.

1.1. Numerical Examples (i.e., Payoff Conditions)

Six numerical examples of the current quantification of Heinrich's safety problems are examined in this paper. In addition to the three characteristics described above, the selected examples satisfy the following constraints (that are likely to hold in real life safety situations): (1) a safe act is costly (at least in a relative way). For example, choosing to drive slowly or to hold back production has some cost in terms of time or money. (2) The risky act can result in heavy loss (in the case of an accident) or in some gain (if an accident does not occur). (3) These characteristics imply that the value of Hit (i.e., a safe act in an unsafe environment) is lower than the value of CR (a risky act in a safe environment), and that the cost of FA (an unnecessary safe act when the environment is safe) is lower than the cost of Accident (as a subresult of Miss). (4) An important property of safety is that the absence of an accident is a deficient feedback for the choice of a safe act since it does not reveal the state of the environment; i.e., one can never be sure if the safe act was really necessary (Reason, 1995). Reason's point implies that Hit and FA outcomes share identical values. (5) In the same manner one can argue that the absence of an accident has the same deficient effect when an unsafe act has been taken and that CR

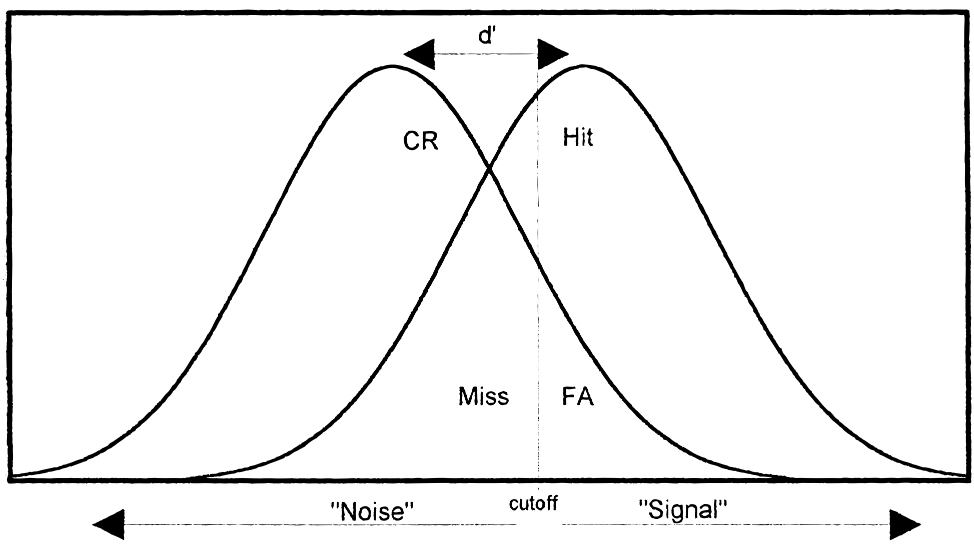

FIG. 1. The Signal and Noise distributions in signal detection theory. The distance between the means of the two distributions (d') is the sensitivity. The cutoff point above which the decision maker indicates that this is a signal specifies the response criterion β.

and Near (the second subresult of Miss) outcomes will also have identical values.

Thus, the rank order of the possible outcomes shown in Table 1 is summarized as follows:

$$v(\text{Accident}) < (v(\text{Hit}) = v(\text{FA})) < (v(\text{CR}) = v(\text{Near}).$$

Six payoff matrices in line with this ordering are presented in Table 2. Whereas these payoff matrices can be used to represent many types of safety problems, it is convenient to consider a specific "cover story." Returning to the chemical plant example, assume that an operator observes a signal indicating that the chemical container may be overfilled and has to decide whether to empty it or ignore the risk and start the heating process. The payoff matrix in Table 2a represents a hypothetical decision maker who perceives the safe act (emptying the container) as costless; i.e. for this decision maker taking the safe action causes no loss or gain independent of the state of nature. A risky act (starting the heating process) yields a small gain (if the state is safe) or a costly accident (if the state is unsafe). (Note that in this example the gamble has only one outcome with the probability $P = 1$.)

TABLE 2

Six Payoff Conditions Representing Heinrich's Safety Problem

a

	State of environment	
Decision	Unsafe	Safe
Safe act	0	0
Risky act	$\begin{Bmatrix}(-10, 1)\\(1, 0)\end{Bmatrix}$	1

b

	State of environment	
Decision	Unsafe	Safe
Safe act	−1	−1
Risky act	$\begin{Bmatrix}(-11, 1)\\(0, 0)\end{Bmatrix}$	0

c

	State of environment	
Decision	Unsafe	Safe
Safe act	0	0
Risky act	$\begin{Bmatrix}(-21, .5)\\(1, .5)\end{Bmatrix}$	1

d

	State of environment	
Decision	Unsafe	Safe
Safe act	−1	−1
Risky act	$\begin{Bmatrix}(-22, .5)\\(0, .5)\end{Bmatrix}$	0

e

	State of environment	
Decision	Unsafe	Safe
Safe act	0	0
Risky act	$\begin{Bmatrix}(-109, .1)\\(1, .9)\end{Bmatrix}$	1

f

	State of environment	
Decision	Unsafe	Safe
Safe act	−1	−1
Risky act	$\begin{Bmatrix}(-110, .1)\\(0, .9)\end{Bmatrix}$	0

While the payoff matrix in Table 2a allows for gain (referred hereafter as a frame of gain), no such gain is possible in Table 2b (referred hereafter as a frame of loss). This payoff matrix represents a second hypothetical decision maker that, unlike the first, perceives the safe act as costly (in absolute terms; for example, loss of time). When this decision maker diagnoses a safe state correctly (CR) he or she neither wins nor loses since it is his/her custom to start the heating process without a second thought. If he or she takes a risky act given an unsafe environment (i.e., Miss), an accident occurs with probability 1. Note that Table 2b was created by a subtraction of a constant (1 unit) from each cell in Table 2a. This was done to ensure an identical expected value maximization strategy under the two payoff matrices. We will utilize this point in the following sections.

The first two examples follow traditional SDT (the gamble was reduced to a deterministic outcome since P equaled 1). The remaining four payoff matrices relax this assumption and allow for a probabilistic penalty for Miss errors. Two of them are directly derived from the payoff matrix in Table 2a (a frame of gain), and the other two are derived from the payoff matrix in Table 2b (a frame of loss). The only change is in the Miss (gamble) cell (the value of the penalty is higher while the probability for that penalty is lower, keeping the expected value of Miss constant). In these four payoff matrices, an error of the decision maker can go unpunished. In Heinrich's terms, failure to detect an unsafe state (or a conjunction between an unsafe state and a risky act) does not always result in an accident. In the payoff matrices presented in Tables 2c and 2d the probability for penalty under Miss is $P = .5$; i.e., only half of the Miss errors lead to an accident. In the payoff matrices presented in Tables 2e and 2f the probability for penalty under Miss is $P = .1$; i.e., only one tenth of Miss errors lead to an accident. The six numerical examples can be summarized into a factorial design of frame $\times$ probabilistic penalty. Three examples represent a frame of gain (i.e., the safe act is costless), while the other three represent a frame of loss (i.e., the safe act is costly). For each frame there are three different levels of probabilistic penalty for Miss ($P = 1$, .5, or .1).

2. BEHAVIORAL MODELS

Human behavior in these cases can be predicted by several quantitative decision-making models. Four alternative models representing four different approaches to risk taking will be considered here. To facilitate derivation of the predictions of all the models we consider for the six examples described above, we add two additional numerical assumptions: $d' = 1.5$ and the prior probability p(Unsafe state) = .3. These assumptions were utilized in the experiment described below.

SDT Ideal Observer Model

Since we are dealing with a signal detection task, the SDT's ideal observer model is a natural choice for predicting behavior. According to traditional

SDT an observer is assumed to use a cutoff decision rule, meaning that a unidimensional stimulus will be labeled “Signal” if its intensity equals or exceeds the cutoff point (referred to as β) and otherwise labeled as “Noise” (see Fig. 1). In order to maximize expected outcome the observer has to set an optimal cutoff (referrer to as β*). Under the assumption of risk neutrality, calculation of the optimal strategy is not affected by the probabilistic nature of the Miss outcome. Rather, only the expected values of the different outcomes are important. The optimal cutoff is sensitive both to the four outcome values and to the prior probabilities of Signal and Noise (P(S), P(N)). Application of this logic to the current task implies that the probabilistic Miss outcome can be replaced by its expected value. Thus, the optimal cutoff is given by:

$$\beta^* = P\,(x|\mathrm{S})/P\,(x|\mathrm{N}) = (\mathrm{CR} - \mathrm{FA})/(\mathrm{Hit} - \mathrm{EV(Miss)}) * (P(\mathrm{N})/P(\mathrm{S})).$$

The ideal observer model predicts that an observer will act *as if* he or she computes the optimal cutoff β*. All six numerical examples (payoff conditions) were designed to have identical risk neutral (RN) optimal cutoffs ($\beta^* = 0.233$, $\ln \beta^* = -1.457$). The prediction of the SDT ideal (RN) observer is therefore that behavior, reflected in cutoff placement, should be optimal and identical for all six conditions. According to this model, behavior would not be affected by probabilistic penalty nor by the different frames (gain or loss). In addition, under the ideal observer model the estimated diagnostic sensitivity (d') is independent of the payoff matrix. It is determined by the distance (in standard deviation units) between the means of the two distributions (see Fig. 1), and should be 1.5 in the current examples.

Note, however, that calculation of the ideal observer’s predictions utilizes information concerning the payoff matrix and prior probabilities. It is unlikely that in industrial settings of the type studied by Heinrich this information is explicitly available, and it will not be provided in the experiment described below, so that the ideal observer does not have a clear prediction for the current setting. However, the prediction can be adjusted to the current setting under the assumption that the observer will learn the necessary parameters from the available feedback and will arrive at the optimal prediction through experience.

An Adaptation of Prospect Theory—the PT Observer Model

It can be demonstrated that the ideal observer model is a corollary of the expected value (EV) rule (see Sperling & Dosher, 1986). Thus, observations that expected utility theory is often violated (Kahneman & Tversky, 1979) suggest that the model has to be modified in order to describe choice behavior. Following Kahneman and Tversky (1979) the current model assumes that the observer tries to maximize the S value function assumed by prospect theory (i.e., concave for gain, convex and steeper for loss). In this model, referred to as the “PT observer,” choice is assumed to be characterized by risk aversion in the gain domain and by risk seeking in the loss domain.[1] Recall that β^* is calculated under the assumption of risk neutrality. In the PT observer model, β is expected to fall below β^* (i.e., below .233) for gain, and above .233 for loss. A weaker prediction of this model is that the observer's cutoff will be relatively lower for gain as compared to a higher cutoff for loss.

Hill-Climbing Model

The hill-climbing model was developed by Busemeyer and Myung (1992) to describe learning processes in SDT tasks.[2] According to the HC model an observer performing a signal detection task modifies performance by a step-to-step shifting of the cutoff. For each step of n trials, the HC observer uses a specific cutoff. To determine the cutoff for the next step the observer compares the average payoffs of the two previous steps. Shifting will be made in the direction of the cutoff which resulted in the higher average payoff. For example, if each step constitutes one trial, ($n = 1$), the HC observer is assumed to reflect upon the two preceding trials when facing trial t. If the last cutoff shift from $t - 2$ to $t - 1$ resulted in a larger payoff, the cutoff will shift on trial t in the same direction. If the last shift resulted in a lower payoff, the next cutoff will shift in the opposite direction. If the last cutoff shift made no difference, the HC observer will choose the next shift at random, as if to check again his position. The model has a "shift size" (adjustment) parameter that implies a finite number of cutoffs. Busemeyer and Myung's quantification is presented in Appendix A. The HC model can be considered as a variant of directional learning (Selten, 1996) where step size n determines the pace of learning. An elegant experimental design (Busemeyer & Myung, 1992) showed hill-climbing predictions to be superior to such classical models as the ideal observer and common error correction models (reviewed by Kubovy & Healy, 1977).

For the payoff conditions suggested above, the HC model predictions are sensitive to the model's step size (the free parameter n). Given a relatively small step size, it predicts risky behavior. There are two reasons for this prediction. First, the majority of the stimuli belong to the Noise category, so the "Noise" response in these cases results in higher values (CR) than the "Signal" response (FA). Second, when the stimulus belongs to the Signal category there is still a chance that the "Noise" response will go unpunished (a near accident may even result in a positive value). Given larger step sizes the HC model predicts less risky choices since average payoffs of larger steps serve as better estimates of the EV for a specific cutoff. With better estimates (i.e., very large

[1]This adaptation of prospect theory is only partial. It is focused on the value function suggested by Kahneman and Tversky (1979), yet it does not take into account the decision weights which are another important component of the theory.

[2]The hill-climbing model is a submodel developed by Busemeyer and Myung (1992) as part of a larger model of decision rule learning which is composed of two parts: an adaptive network model that determines the decision rule for a given situation, followed by the hill-climbing process for fine tuning the chosen rule's parameters. We concentrated on the hill-climbing model because the decision rule used for SDT tasks is a cutoff rule.

step sizes) this model converges to the ideal observer prediction. Probabilistic penalty causes further damage to these estimates resulting in the prediction of riskier cutoffs (i.e., higher) as the probability for penalty decreases. Finally, the HC model predicts no framing effect since it is more sensitive to the ranking order than to the exact quantity of outcomes.

To derive the hill-climbing model predictions of the estimated cutoff placement (β) and diagnostic sensitivity (d') for the six payoff conditions, we ran two sets of computer simulations, using the original parameters of Busemeyer and Myung, with a small step size ($n = 1$) for one set and a large step size ($n = 10$) for the other set. The remaining parameters were: the most extreme cutoff $C_{\min} = -5$ (in standardized units) and number of cutoffs $M = 41$, the magnitude of change = .25, the probability of an increase (and a decrease) given identical payoff in the last two trials = .5 (see Busemeyer & Myung, 1992, for raw parameters and additional details). Specific predictions for ln β and for d' are presented in the third column of Figs. 4 and 5 and are discussed later.

Cutoff Reinforcement Learning Model

The last model is the cutoff reinforcement learning model proposed by Erev, Gopher, Itkin, and Greenshpan (1995). The CRL model assumes that the observer's choices among possible cutoffs can be described by the law of effect (Thorndike, 1898) as quantified by Roth and Erev (1995). This states that the probability of adopting a certain strategy (i.e., a cutoff) increases if this strategy is positively reinforced and decreases if it results in negative outcome. According to the CRL model, each observer has an initial propensity to choose each of a final set of cutoffs. The probability of selecting cutoff c is determined by the ratio of its specific propensity and the sum of all other existing propensities. Positive reinforcement increases the propensity of the selected cutoff, while negative payoff reduces it. Erev *et al.*'s (1995) quantification is presented in Appendix B. Unlike the HC model, which comprises a systematic search for the best cutoff, the CRL model allows for random experimentation with different cutoffs over a wide range. This may slow the learning rate, but it may also help to avoid local maximum or suboptimal cutoffs. Erev (1998) demonstrated that the CRL model can account for robust violations of the ideal observer model. Among the 19 violations considered by Erev (1998) are conservatism and probability matching. He also presented a comparison between the CRL and the HC models, showing that in the tasks studied by Busemeyer and Myung (1992) the two models have similar descriptive values.

The CRL model was recently shown to account for behavior in safety-like SDT tasks. Gopher, Itkin, Erev, Meyer, and Armoni (1995) showed that predictions of the CRL model held for both implicit and explicit cutoff paradigms and for one-person tasks and two-person shared responsibility tasks. The payoff matrix used in these studies was identical to the one shown in Table 2a. The prediction of CRL for this payoff matrix (in a one-person signal detection task) included an initial trend of risky behavior ("learning away" from the optimal cutoff in the short run) followed by a consistent trend to lower cutoffs toward the optimal one. Yet, the CRL model was not applied to probabilistic signal detection tasks as in the present work. The meaning of probabilistic penalty in the CRL model's terms is that risky cutoffs might be reinforced even when they lead to a Miss. Reinforcement of risky strategies will increase as the probability for penalty decreases.

To derive the predictions of the CRL model for the six payoff conditions, computer simulations were run using the original parameters of Erev *et al.* (1995). These include two strategy space parameters (the most extreme cutoff $C_{\min} = -5$ and number of cutoffs $M = 101$)[3]; two initial propensity parameters (initial standard deviation $\sigma_i = 1.5$, and initial strength $s(1) = 3x$ (the average payoff from random choice)); three reference point parameters (an initial reference point $\rho(1) = 0$ and weights for positive and negative outcome $w^+ = .01$, $w^- = .02$); a generalization standard deviation $\sigma_g = .025$; a recency parameter $\varphi = .001$; and a technical parameter to ensure that all propensities are positive ($v = .0001$). The model predictions for ln β and for d' are presented in the fourth column of Figs. 4 and 5, and are discussed in the Results section below.

In summary, inspection of Figs. 4 and 5 reveals that each of the four models yields different predictions for the examples of Heinrich's safety problem considered here. The ideal observer predicts convergence to optimal and identical behavior for all six conditions. The PT observer model predicts a framing effect: risk aversion for gain matrices (Tables 2a, 2c, 2e) and risk seeking for loss matrices (Tables 2b, 2d, 2f). The HC model predicts stable risky behavior. This prediction is more extreme for the small step size and moderate for the large step size. Finally, the CRL model predicts a pattern of initial tendency toward risky behavior, followed by a gradual shift to safer behavior. Both the HC and CRL models predict that behavior will become riskier as the probability for penalty is lowered. A signal detection experiment, designed to compare the four alternative sets of predictions, is presented in the next section.

3. EXPERIMENT

In order to compare the four models, the experiment examined choice behavior under the six numerical examples (payoff conditions) for which the models' predictions were derived. The choice of the six conditions implies that, in addition to model comparisons, the experiment will shed light on some "model free" questions, i.e., the effects of probabilistic penalty, and gain and loss frames, in probabilistic signal detection tasks. The experimental task was adapted from Lee (1963), displaying different heights arising from either of two experi-

[3]Using the original parameters of CRL and HC models resulted in an unequal number of cutoffs. While the CRL model defines explicitly the set of 101 cutoffs, the HC model's set of cutoffs is determined by the magnitude of change in every shift. We used the original magnitude of change (in std units) which defined 41 cutoffs. Using more cutoffs for the HC model would have caused it to predict slower learning.

mentally determined normal distributions. Administrated d' (the distance between the means of the two distributions) was 1.5 standard deviation units. Signal (an unsafe state) was represented by the higher distribution with $p(S) = .3$ while Noise (a safe state) was represented by the lower distribution and was given with the complementary $p(N) = .7$ (see Fig. 2 for illustration). The six payoff conditions imply a 2×3 factorial design of frame-type $\times$ probabilistic penalty. Three statistical indices of interest were Miss rates, the representative of potential accidents; β, the measure of cutoff placement which represents the strategic tendency toward safe or risky behavior; and d', the measure of the diagnostic sensitivity of subjects.

3.1. Method

Participants. Ninety Technion students participated in the experiment. They were recruited by advertisements offering monetary reward of 20–35 NIS ($6.67–$11.67) according to their performance during a 1-h computerized experiment. Average payoff was 25 NIS ($8.33).

Apparatus. The experiment was programmed using Visual Basic 3 for Windows on a 486PC with a Super VGA 14″ screen. Figure 2 presents the stimuli shown on the screen and illustrates the underlying distributions which were not shown to participants. A green rectangle 14 cm high and 7.5 cm wide was placed in the center of a 19 × 26 cm blue background. Two feedback fields

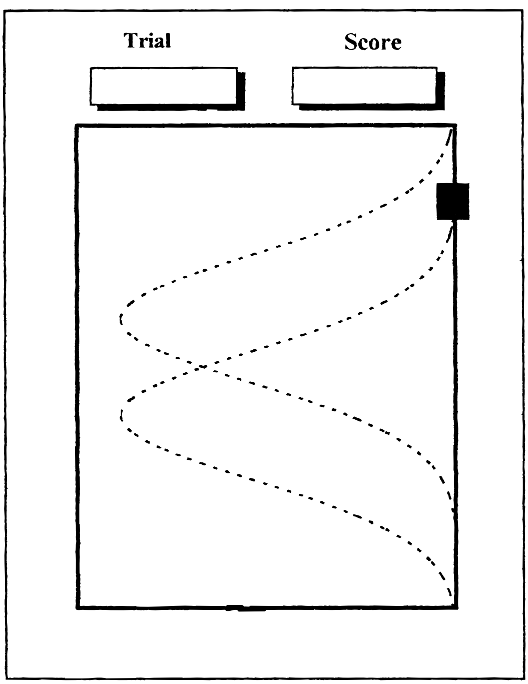

FIG. 2. A schematic depiction of the experimental task. The stimulus with underlying distributions. Distributions are not shown to the subjects and are only added here for demonstration purposes. Feedback for last trial and cumulative scores is presented in the two text boxes.

were located 1 cm above the rectangle (white text boxes of 1×2 cm). The right one was labeled "score" and the left one was labeled "last trial." A white 2×2 mm square at different points along the right border of the rectangle served as stimulus and represented different heights relative to the rectangle. The position of the square was determined by a two stage sampling—one for selecting High or Low distribution and the other for selecting 1 out of 66 possible locations according to the selected normal distribution.

Each trial began with a 2-s presentation of the white square. Response was made by pressing one of two keys on the keyboard (⟨3⟩ for "low" and ⟨6⟩ for "high" on the right number panel). After the response was made, feedback was given as to the trial score and the cumulative score (i.e., "last trial" and "score"; see Fig. 2). The trial score was given according to the relevant payoff matrix, with random choice between penalty or no penalty for Miss. Cumulative score was adjusted accordingly. Feedback was presented for 3 s before beginning the next trial.

Procedure. Participants received written instructions which were also read aloud. They were told that they are playing a game in which they have to decide whether a presented white square is high or low. They were given a hint about the probabilistic nature of the square using an example taken from Kubovy and Healy (1977) about categorizing a person as a man or a woman according to his/her height. The experiment consisted of six blocks of 100 trials each. At the beginning of each block participants received 2000 points. In each trial they could earn or lose points according to their decision and according to the payoff matrix employed (with the loss matrices participants were aware that they should minimize losses). A block was terminated with an on screen announcement "End of Block," and another block started when the participant pressed the "Enter" key. Participants were told that there were six blocks for the game and that their goal was to win as many points as they could in every block. Participants were not informed about the payoff matrix. Instead, they were told that they have to learn the possible outcomes from the step-to-step feedback. Nor did participants know the exact monetary value of each point, but that payment would be proportional to the score and is given according to a random choice of one of the six scores they achieve in the entire game.[4] To fully understand the task, participants were given 2 min training with a symmetric payoff matrix. This also served to monitor initial propensities and to direct initial strategies to the natural startup point $\beta = 1$. After training, subjects were told that the experiment was about to begin and that points would now be differently allotted, which they would have to learn through the step-by-step feedback.

[4]Roth and Maluf (1979) noted that when the specific relation of monetary payoff to "points" is probabilistic, the expected utility theory implies the EV rule. We did not inform subjects of the details of the payoff rule in advance (they only knew that the probability of a bonus increases with the number of points), as experimental results (Selten *et al.*, 1996) suggest that this may lead to confusion. Payment was proportional to the points with a step of minimal payment: The initial 2000 points were equivalent to a minimal payoff of 20 NIS. Each extra point added .2 NIS.

Experimental design. As noted above, the six payoff conditions resulted in a 2×3 between subjects design. Three different levels of probabilistic penalty were used in the frame of gain and three in the frame of loss (see Tables 2a–2f). Learning was measured within subjects through repeated measurements on six blocks.

3.2. Results

Descriptive Statistics

Table 3 presents the main experimental results by experimental condition and block. Using the common signal detection statistics, Table 3 displays the mean rate of Hit (p(Hit)), False-Alarm (p(FA)) and Miss (p(Miss)). Also presented are the means of the derived statistics, ln β and estimated d' (which represent the strategic cutoff placement and sensitivity, respectively).[5] We use the transformation from β to ln β in order to avoid the skewed distribution of β.

Model-Free Hypothesis Testing

Rate of Miss errors. Figure 3 presents the rates of Miss errors for each experimental condition by blocks. Errors of this type were the focus of Heinrich's study because they can lead to accidents. Table 3 and Fig. 3 indicate that the miss rate increases with a decrease in the probability of being penalized. A decline in the miss rate was observed over time. A three-way ANOVA (probabilistic penalty $\times$ frame $\times$ block) with repeated measures on block was conducted to evaluate the significance of these trends. The analysis reveals a significant main effect for probabilistic penalty, indicating a consistent increase in Miss errors as the probability for penalty decreases ($F[2,84] = 88.28$ $p < .0001$). The effect of learning (block) was also significant indicating that experience created a consistent decrease in Miss errors ($F[5,420] = 26.67$ $p < .0001$). The framing effect was significant only at the lowest level of probabilistic penalty (.1) where the number of errors was higher for the frame of loss than for gain ($F[1,84] = 10.14$, $p = .002$).

Response criterion ln β. The right-hand column in Figs. 4a–4c presents graphically the results for ln β. The models' predictions are given to the left of the empirical data. The HC and CRL models' predictions for each experimental condition show the average of 1500 simulations (100 replications, each with 15 virtual subjects) summarized (like the experimental data) in six blocks of 100 rounds. Recall that higher values reflect risky cutoffs and lower values reflect cautious cutoffs.

[5]SDT statistics cannot be calculated given p(Hit) and p(FA) = 0 or 1. To avoid this problem, zeros were replaced by .025, and ones were replaced by .975. Of the 90 (subjects) $\times$ 6 (blocks) = 540 assessments, replacements were needed for cases of zeros for p(Hit) 6 times, for cases of zeros for p(FA) 34 times (29 of them in the .1 loss frame), for cases of ones for p(Hit) 15 times, and for the cases of ones for p(FA) 2 times. Over all, there were 57 replacements.

TABLE 3

Mean Results for Hit, False Alarm, and Miss Rates. Mean Results and (Standard Deviations) for Sensitivity (d′) and ln β

		Frame of gain						Frame of loss					
		Block						Block					
		1	2	3	4	5	6	1	2	3	4	5	6
P (Penalty|Miss) = 1	*P*(Hit)	0.72	0.77	0.81	0.84	0.83	0.84	0.62	0.75	0.81	0.82	0.81	0.85
	P(FA)	0.28	0.31	0.34	0.37	0.33	0.39	0.27	0.38	0.39	0.44	0.43	0.46
	P(Miss)	0.28	0.23	0.19	0.16	0.17	0.16	0.38	0.25	0.19	0.18	0.19	0.15
	d'	1.25	1.39	1.45	1.40	1.51	1.38	1.00	1.18	1.34	1.26	1.23	1.36
		(0.38)	(0.45)	(0.31)	(0.39)	(0.36)	(0.30)	(0.28)	(0.51)	(0.37)	(0.32)	(0.26)	(0.32)
	ln β	0.01	−0.23	−0.46	−0.49	−0.40	−0.50	0.12	−0.25	−0.49	−0.48	−0.49	−0.70
		(0.50)	(0.68)	(0.75)	(0.50)	(0.68)	(0.53)	(0.41)	(0.83)	(0.83)	(0.80)	(0.78)	(0.83)
P (Penalty|Miss) = 0.5	*P*(Hit)	0.53	0.58	0.54	0.68	0.67	0.74	0.48	0.53	0.53	0.65	0.70	0.79
	P(FA)	0.21	0.19	0.20	0.30	0.32	0.38	0.20	0.25	0.28	0.30	0.35	0.44
	P(Miss)	0.47	0.42	0.46	0.32	0.33	0.26	0.52	0.47	0.47	0.35	0.30	0.21
	d'	0.99	1.31	1.13	1.22	1.17	1.25	0.89	0.89	0.81	1.05	1.07	1.08
		(0.39)	(0.39)	(0.47)	(0.36)	(0.41)	(0.51)	(0.37)	(0.61)	(0.61)	(0.63)	(0.59)	(0.57)
	ln β	0.40	0.52	0.44	0.01	−0.02	−0.18	0.46	0.20	0.24	−0.05	−0.04	−0.39
		(0.59)	(0.90)	(0.80)	(0.85)	(0.99)	(1.03)	(0.45)	(0.75)	(0.66)	(0.67)	(0.70)	(0.69)
P (Penalty|Miss) = 0.1	*P*(Hit)	0.38	0.40	0.43	0.57	0.63	0.66	0.22	0.25	0.37	0.33	0.46	0.40
	P(FA)	0.17	0.20	0.21	0.26	0.35	0.36	0.05	0.10	0.16	0.19	0.22	0.19
	P(Miss)	0.62	0.60	0.57	0.43	0.37	0.34	0.78	0.75	0.63	0.67	0.54	0.60
	d'	0.74	0.76	0.63	0.96	0.88	0.94	0.85	0.68	0.72	0.49	0.85	0.76
		(0.53)	(0.76)	(0.89)	(0.66)	(0.67)	(0.78)	(0.37)	(0.52)	(0.52)	(0.67)	(0.52)	(0.96)
	ln β	0.43	0.49	0.19	0.29	−0.01	−0.10	0.96	0.63	0.38	0.23	0.32	0.39
		(0.40)	(0.92)	(0.44)	(0.46)	(0.43)	(0.27)	(0.40)	(0.46)	(0.35)	(0.33)	(0.48)	(0.39)

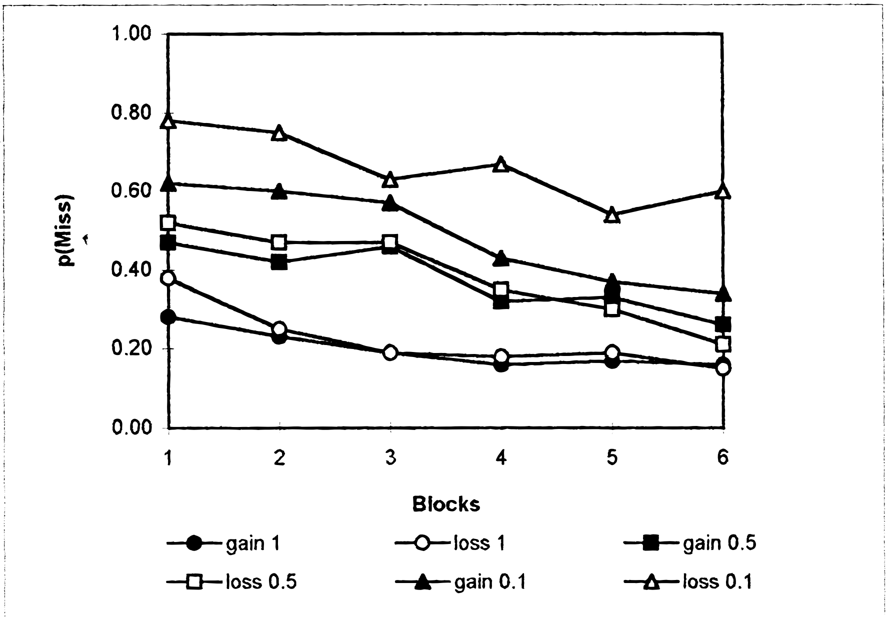

FIG. 3. Miss rates (representing critical errors) across payoff conditions. The miss rate is larger with small penalty probabilities.

A three-way ANOVA revealed a significant main effect of probabilistic penalty ($F[2,84] = 16.32$, $p < .0001$) maintaining increased risk taking with decreased probability of penalty. The main effect of learning (block) is also significant ($F[5,420] = 22.16$, $p < .0001$) and it indicates that experience leads to a consistent decrease in ln β values toward the optimal point. The main effect of the frame was not significant ($F[1,84] = .13$, ns) indicating that, in general, subjects neither took riskier choices under the frame of loss nor behaved more cautiously or avoided risk in the frame of gain. Even so, a short-run relative framing effect was significant in the first block ($F[1,84] = 5.94$, $p < .017$). Notice that at all three levels of probabilistic penalty, the first block shows riskier cutoff for loss compared to gain. Closer inspection indicates that the statistical source of the short-run framing effect was the first block of the .1 probabilistic penalty, in which the cutoff for the loss matrix was markedly riskier than the cutoff for the gain matrix ($F[1,84] = 9.87$, $p < .0023$).

Sensitivity d'. The right-hand columns in Figs. 5a–5c present graphically the results for d'. The three alternative predictions for d' are presented graphically to the left of the empirical data (the HC and CRL models predictions show the average of 1500 simulations). Analysis of variance shows a main effect for probabilistic penalty ($F[2,84] = 15.02$, $p < .0001$) indicating that estimated d' decreases with a decrease in the probability of penalty. The effect of learning was also significant, as the values of d' increased with practice ($F[5,420] = 2.45$, $p < 0.0332$). Interestingly, unlike ln β, d' is apparently influenced by the frame. A significant framing effect ($F[1,84] = 3.93$, $p < 0.05$) implying that d' values were lower in the loss domain was observed. The loss of sensitivity in the frame of loss is especially clear in Figs. 5a and 5b (at the

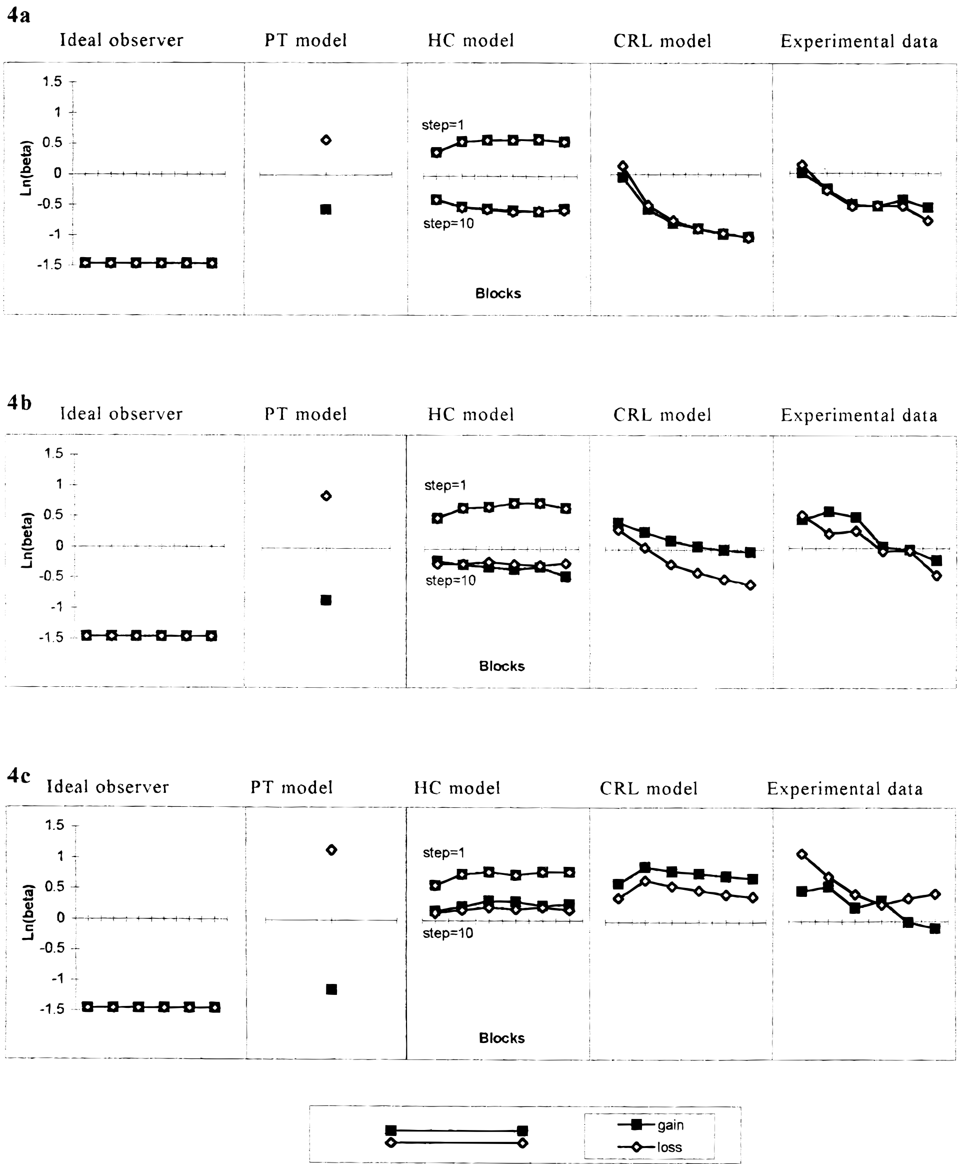

FIG. 4. Average ln β values across payoff conditions. The filled squares represent ln β for the gain matrices, and the empty diamonds represent ln β for the loss matrices. Predictions are shown from left to right (1) SDT ideal observer, (2) PT observer, (3) Hill-climbing model (higher values for step size = 1 and lower values for step size = 10), and (4) CRL model. Experimental results are presented in the right-hand column.

lowest level of probabilistic penalty *d′* values are very low for both frames; see Fig. 5c).

Model Comparison

Qualitative comparison for response criterion (ln β). A comparison of the models' predictions and the actual results is provided in Fig. 4. As can be seen,

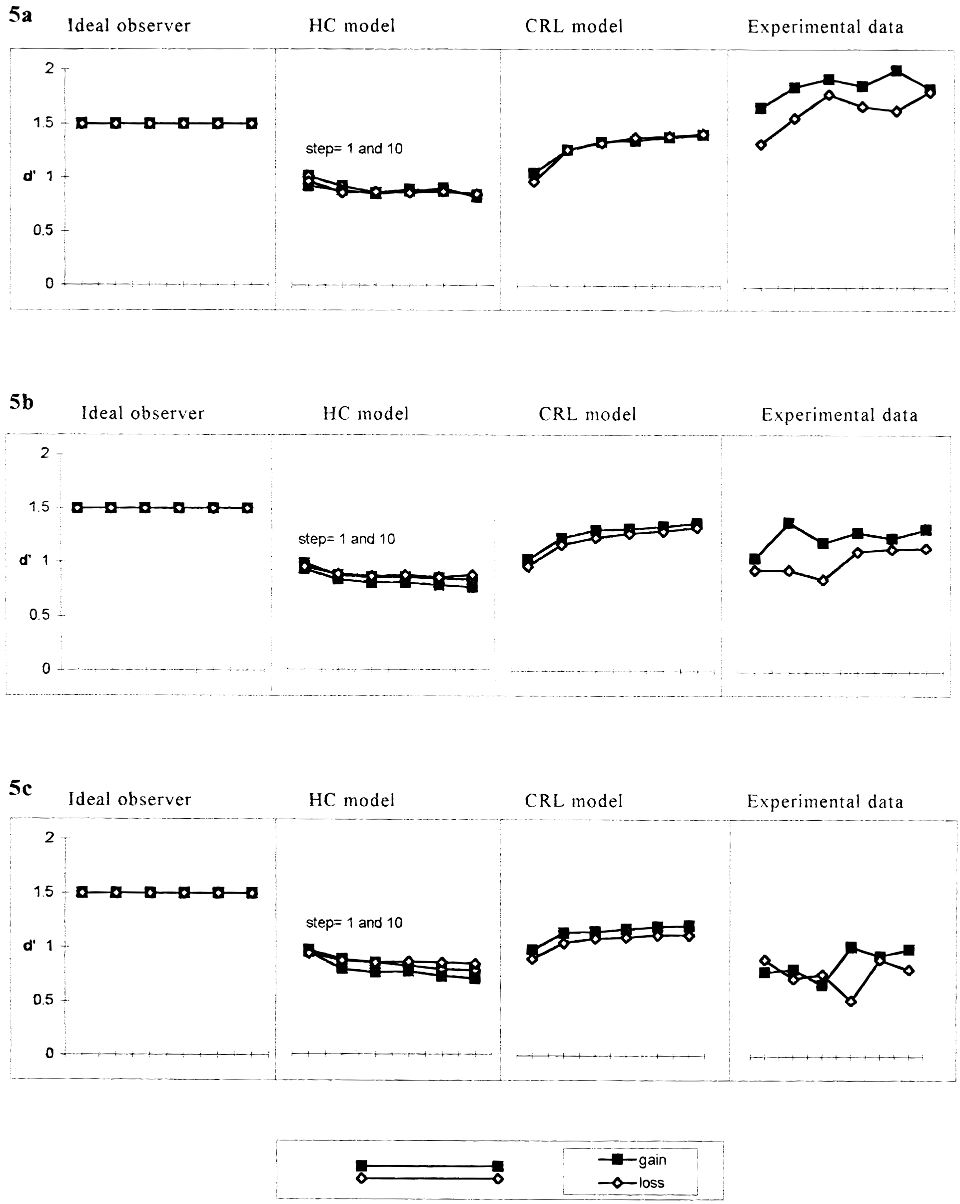

FIG. 5. Average d' values across payoff conditions. The filled squares represent d' for the gain matrices, and the empty diamonds represent d' for the loss matrices. Predictions are shown from left to right (1) SDT ideal observer and PT observer, (2) Hill-climbing model (small and large step sizes yield similar predictions), and (3) CRL model. Experimental results are presented in the right-hand column.

the two static models do not provide a good approximation for behavior. In addition to overlooking the significant learning trend, these models do not capture the behavior of experienced players. Specifically, subjects took greater risks than predicted by the ideal observer model and were sensitive to the probability of penalty which was not predicted by this model. The PT observer model captures the qualitative trends in the first block but fails to predict learning toward the optimal cutoff.

The two dynamic models were closer to actual behavior. The HC model predictions served as the upper and lower bounds of the experimental results. Initial risky behavior was in agreement with the smaller step size prediction, and closing (safer) behavior was in agreement with the larger step size prediction. Both sets of the HC predictions were in agreement with the main effect of probabilistic penalty. The CRL model also captured the effect of probabilistic penalty. In addition this model captured the initial risk taking behavior that represents learning away from optimal behavior. In line with the prediction of this model, the counterproductive risk taking was larger with small penalty probabilities. Finally, note that neither CRL nor HC models captured the short-run framing effect.

Qualitative comparison for sensitivity d'. In line with the cutoff statistics the dynamic models provide better predictions of sensitivity (d' values as presented in Figs. 5a–5c). Contrary to the static models' predictions for d' (i.e., $d' = 1.5$), the experimental results showed a gradual increase with time, as well as probability and frame effects. The HC model captured the decline in d' values due to a probabilistic penalty, but not the increase in d' with practice. The effect of a probabilistic penalty on d' was also captured by the CRL model which, in addition, captured the effect of practice on d' and the experimental framing effect, i.e., higher d' values for the gain and lower values for loss.

Note that in simulations that were run to derive the predictions of the CRL and the HC models, the virtual subjects had true (virtual) d' of 1.5 (i.e., the distance between the two observed distributions was 1.5 STD units). The estimated d' values in these simulations were lower than 1.5. This was a result of the continuous cutoff changes in the HC model and of the probabilistic choice rule in the CRL model. The resulting inconsistency reduces the estimated d'. Thus, the relatively good fit between the experimental d' and the dynamic models' predictions suggests that the low experimental d' might reflect response variability rather than perceptual ability.

Quantitative model comparisons. To allow for quantitative comparison of the two dynamic learning models we created a data set that includes the observed results and different predictions for each of the 36 experimental blocks (6 conditions $\times$ 6 blocks) for each statistic ($\ln \beta$ and d'). Table 4 presents two types of quantitative fitness scores derived from these data. These include the correlations between the experimental data and model predictions, and MSD scores (mean squared differences between model predictions and the experimental data).[6]

Both fitness scores reveal a consistent order in which the HC model with step size 1 has the lowest fitness scores (for both $\ln \beta$ and $d)'$, the HC model with step size 10 has better scores, and the CRL model offers the best scores.

[6] The F statistics for the MSD scores test whether the experimental data diverge significantly from the predicted means.

TABLE 4

Dynamic Learning Models' Fitness Scores: Pearson Correlation (r(model,experimental)), MSD Estimates, and Related *F* Statistics

		Statistic	
Fitness score	Model	ln β	d′
Correlation	HC (step = 1)	.25	.04
	HC (step = 10)	.71**	−.01
	CRL	.83**	.69**
MSD	HC (step = 1)	56.95 $F[4,30] = 7.30$*	16.19 $F[4,30] = 3.22$*
	HC (step = 10)	16.14 $F[4,30] = 2.07$ ns	17.78 $F[4,30] = 3.54$*
	CRL	13.40 $F[4,30] = 1.72$ ns	6.33 $F[4,30] = 1.26$ ns

* $P < .05$.
** $P < .01$.

Positive correlations between the models' mean predictions and the mean results for ln β were observed in all three cases, so that all sets of predictions included the trend of response criterion shifts with practice. The correlations for the HC model with step size 10 and the CRL model reached significance. Only the CRL predictions were significantly correlated with the observed d' scores. The HC model predictions (i.e., a decrease in d' with practice) were inconsistent with the experimental d' values. MSD statistics agree with the correlations pattern (see Table 4).

These results suggest that whereas both the HC and the CRL models capture important aspects of the data, the CRL model provided better qualitative and quantitative approximation. However, some important questions remain to be answered. How do different search rules arrive at similar predictions? Why does the HC model fail to capture the effect of experience on d'? Why does the PT observer model fail to capture behavior in the current task? We address these and other points in the next section.

4. DISCUSSION

The present research demonstrates that Heinrich's (1931) characterization of the sources of industrial accidents can be modeled as a probabilistic signal detection task. This abstraction facilitates the derivation of quantitative predictions and the experimental examination of behavior. The experimental results, summarized above, provide indications of six behavioral tendencies in these tasks: (1) initial learning toward suboptimal risky behavior; (2) slow learning toward safer optimal cutoff placement over the long term; (3) decreased learning speed (and increased miss rate) with decreased probability of penalty; (4) slow increase in d' with experience; (5) d' impaired by probabilistic penalty; and (6) d' impaired by the absence of positive outcomes. These results suggest that in the tasks under study, dynamic models provide a better approximation of behavior than static models. In addition, the results suggest that small quantitative details can have a robust effect on initial risk taking behavior and on the speed of adaptation/learning process. It appears, then, that quantification of Heinrich's observation, and of the dynamic decision process, can lead to interesting implications.

Theoretical Implications

The relative success of the two dynamic models suggests that their common assumptions have a descriptive value. These assumptions state that (1) subjects choose (or behave as if they do) among cutoff strategies and that (2) these choices are sensitive (in an adaptive sense) to the obtained payoff experience. These assumptions are apparently sufficient to capture the main behavioral trends.

The two models differ with respect to the assumed quantitative relation between the obtained payoffs and cutoff selections. Whereas the HC model implies qualitative directional learning, the CRL model assumes a quantitative probabilistic response that is sensitive to accumulated reinforcements. To understand this difference, it should be noted that the HC model violates two basic principles of the CRL model: the law of effect and the power law of practice. The law of effect is violated by the assumption that a reinforced cutoff will be discarded during the systematic directional search. The power law of practice implies stabilization of behavior with practice. This principle is violated by the assumption that the step size (which determines the learning pace) does not depend on the learner's accumulated experience. Stabilization is denied by the continuous change of the cutoff at fixed intervals and magnitude. The significance of the power law of practice is most apparent in the case of the observed d'. An increase in the observed d' with practice reflects the increased consistency of cutoff choice (Kubovy & Healy, 1977). The CRL model predicts this effect because it gradually reduces the number of relevant cutoffs and the magnitude of changes. The HC model fails to predict the observed d' because of the inconsistency inherent to its search rule.

However, this work was not designed to determine which of the four considered theoretical approaches is the "accurate" one. Rather, we compared specific generalizations of these approaches in a specific setting. The results do not rule out the possibility that in the long run, subjects may arrive at the ideal observer predictions or that if the decisions were presented as numerical gambles, a PT pattern would emerge. The HC model could outperform the CRL model in an explicit cutoff paradigm (Busemeyer & Myung's experimental paradigm). Moreover, we have employed here only one version of the HC model in which adjustments of cutoff are made using a constant. The general model presents a broader assumption, stating that adjustment should be a monotone function of the product of the change in cutoff and the change in average payoff. For example, it is possible to suggest quantitative adjustment for the current setting, in which the change of cutoff equals this exact product. This way, a penalty will have a significant effect on the learning process (even if given rarely). A pilot study using computer simulations shows that with this version, the predicted learning process is very similar to the CRL prediction. Yet, this slight change does not solve the consistency problem. To allow the increase in consistency, an implementation of the power law of practice is needed. For example, one could assume that the step size increases with experience, therefore decreasing the rate of cutoff changes and stabilizing behavior. It is possible that this kind of HC model variation will outperform the CRL model in the considered tasks. The current results simply demonstrate that, in the current setting, the CRL model provided a more useful prediction of behavior; it correctly predicted all six qualitative experimental trends and provided a good quantitative fit. Demonstration of the success of a specific approach does not rule out the possibility that other approaches could also be useful.

Some Limitations and Future Extensions

The implementation of both SDT and dynamic models of decision making have some limitations. One apparent limitation of SDT as a model for safety problems concerns the way it represents the environment. The abstraction of the environment as a series of independent categorization tasks is of course an oversimplification. In many industrial settings sequential tasks are correlated. For example, if a system breaks on trial t it is likely to stay broken at $t + 1$ (unless it has been fixed). The first author is now extending the present work using continuous stimuli, investigating this and related questions. Another oversimplification regards the complexity of the signals. Whereas the current experiment focused on unidimensional stimuli, most natural safety problems involve multidimensional stimuli. Under traditional SDT the additional dimensions do not create a problem since they can be reduced to a single dimension. Yet, this assertion is not likely to hold descriptively (Busemeyer & Myung, 1992). Future research is needed to address multidimensional stimuli.

The current implementation of the dynamic models does not take into account prior expectations that are likely to affect behavior in natural settings. For example, think about the detection of car tracks during the selection of a spot to place a sleeping bag. In such a case, a Miss error is more than likely to be fatal. Obviously, most decision makers need no reinforcement learning process to avoid the risky decision in this setting. To take this into account in the CRL model, attention has to be paid to the decision maker's initial propensities. In the above scenario, the decision maker is likely to have a strong inclination against sleeping on the car tracks, which implies that in some cases, choice behavior is sensitive to initial tendencies. Hence, these tendencies have to be precisely modeled. Future research should address this issue. One possible solution could be based on the idea that PT can be used to approximate initial propensities. This solution is supported by the observation of PT's reflection effect in the first block of our experiment.

Another information-related problem of the dynamic models arises in situations in which important, nonpayoff, feedback is provided during the learning period. For example, consider a driver who succeeds in taking an icy turn safely, but is aware of the ice. According to the CRL model, this driver's risk-taking tendency is positively reinforced. However, it is also possible that the driver will consider the new information ("the road is icy") and decide to be more careful. This behavior can be modeled by the CRL model with the assumption that the outcome of "near accident" can be negative reinforcement. Again, this aspect requires future research.

Some Practical Implications

The present work focuses on situations in which the decision makers' feedback can be approximated by their personal "final outcome" experience. We contend that many industrial tasks, as well as daily manual and perceptual tasks, fall into this category. Our experimental results imply that the six behavioral trends observed in the experiment are likely to characterize performance in these tasks. This suggests, for example, that during the initial learning period, the probability of an accident might increase with practice. This can be used when planning optimal training procedures.

Another suggestion involves the effect of positive feedback. We found that such feedback increases consistency and reduces the miss rate. This is consistent with (and provides an explanation for) field results. In a recent paper reviewing 24 safety field studies McAfee and Winn (1989) concluded that safety programs which provided some sort of positive feedback (feedback, verbal prizes, tokens) elicited the desired safety behaviors. The results for probabilistic penalty shed some light on the effect of punishment on unsafe behaviors. The penalty for Miss, which represented the critical error, enabled learning in our payoff conditions. The penalty was the most informative payoff since nature (i.e., a Signal representing danger) was only revealed to the subjects when they were penalized. Using a penalty as feedback for unsafe behavior offers two practical implications. First, the feedback for unsafe behavior should be selective, i.e., punishment should not be given for all unsafe acts, but should offer information regarding specific critical errors. Second, while a penalty may not be excessive, it should be administered for as many errors as possible, in order to arrive at the best learning process.

These implications return us to Heinrich's (1931) recommendation, that emphasis should be given not to the safe or unsafe acts themselves, but to the underlying skills they represent, i.e., diagnosis and decision making. Generalization of this recommendation is offered by cognitive game theory (Erev, 1998; Erev & Roth, in press), which implies that understanding choice behavior is a prerequisite when planning a training program. For this, three factors have to be modeled. First, one has to consider the structural characteristics of the task, such as the incentive structure and the available information. The second is the set of strategies considered by the decision maker. Analysis of the interaction between these two factors will reflect the underlying skills needed for optimal performance. Finally, the effects of different kinds of feedback on the learning process must be comprehended to come up with effective training procedures.

The present research suggests that in many industrial safety problems, the first two factors can be approximated by a probabilistic SDT model (the current quantification of Heinrich's observation). In addition, the current experimental results suggest that the third factor can be approximated by the CRL model. The relative success of the CRL model suggests that it might be used for deriving predictions for other safety problems modeled as probabilistic signal detection tasks. For example, the CRL model can be utilized to predict the effect of the extremely low probability of penalty (i.e., about 1/300) which apparently characterizes industrial accidents (Heinrich, 1931), but cannot be efficiently studied experimentally. Basically, the model predicts that the trends observed here will intensify as probability decreases and that initial learning toward suboptimal risky behavior will take longer. More research in this direction should examine these possibilities.

CONCLUDING REMARKS

Safety problems are being studied from different points of view, aiming at different questions, and using different models and experimental procedures. Obviously, there is no one "best" or "only" way to approach the problems of risk taking in these situations. We suggest that the present work complements the existing research in three ways.

The vast body of the safety research focuses around the themes of causation and intervention. Much of the work regarding causation is very theory oriented (e.g., Petersen, 1988; Reason, 1995). On the other hand the work regarding intervention is more field oriented (e.g., McAfee & Winn, 1989). The present work complements the existing research by utilizing a controlled experiment procedure as a middle stage between theory and field research. In doing so, we were able to investigate the cognitive mechanisms that may both reflect theoretical constructs as well as govern day-to-day behavior in the field. One could argue against the synthetic experimental task we have employed and the lack of its external validity. This shortcoming cannot be ignored, yet our findings are consistent with both theoretical claims and field findings. In the face of this consistency we suggest that it may be safe enough to consider the experimental task as a useful simulation of the decision problem in safety situations. Moreover, this consistency may point out the importance of "ordinary" cognitive processes that underlie behavior in risky situations.

Utilizing signal detection offers two additional advantages for safety research. First, as McKenna (1988) pointed out, qualitative definitions of common concepts such as "uncertainty" "risk," and "human error" are frequently vague and tend to carry on different meanings in different works. The probabilistic SDT model we have suggested here allows clear quantitative definitions for these concepts as well as a specific criterion and quantitative measures of risk taking. Second, the proposed model has also made us able to demonstrate some common assumptions about factors that enhance risk taking, such as the rarity of actual accidents and the imperfect feedback for safe and risky acts. Though these assumptions are widely used in safety research, to the best of our knowledge they have not been tested.

APPENDIX A

The Hill-Climbing Model Learning Rule

Hill-climbing is guided by the change in (average) payoffs, produced by a change in the cutoff from two prior applications of the rule. If the previous change in the cutoff produced an increase in average payoffs (an uphill direction), then the next change is made in the same direction. However, if the previous change in the cutoff produced a decrease in average payoffs (a downhill direction), then the next change is in the opposite direction.

Define θ_1 and θ_2 as the cutoff values that were used on the last and second-to-last steps, respectively. Define v_1 and v_2 as the average values of payoffs produced on the last and second-to-last steps, respectively.

The product of the differences h is called the *hill-climbing adjustment*:

$$h = (\nu_1 - \nu_2) * (\theta_1 - \theta_2).$$

The next cutoff θ is determined to be

$$\theta = \theta_1 + \lambda * m(h) + (1 - \lambda)*\varepsilon,$$

where $m(h)$ is a monotone increasing function of the hill-climbing adjustment h; ε is a randomly chosen direction; and λ is a weight $(0 < \lambda < 1)$ that moderates the amount of random search. Busemeyer and Myung (1992) used a simple example in which $m(h)$ was set as a step function: $m(h) = c$ if $h > 0$, $m(h) = -c$ if $h < 0$, $m(h) =$ random choice of c or $-c$ if $h = 0$, and the effect of the random search was eliminated by setting $\lambda = 1$.

According to this simple example the next cutoff θ is determined to be

$$\theta = \theta_1 + c \text{ if } h > 0; \theta = \theta_1 - c \text{ if } h < 0; \theta = (\theta_1 + c) \text{ or } (\theta_1 - c) \text{ if } h = 0.$$

The HC model was operationalized in this work using the same parameters of this simple example.

APPENDIX B

The Cutoff Reinforcement Learning Rule

Erev *et al.*'s (1995) adaptation of Roth and Erev's (1995) learning rule can be summarized by the following four assumptions:

A finite number of uniformly and symmetrically distributed cutoffs.

A1q. The DM considers a finite set of m cutoffs. The location of cutoff j $(1 \leq j \leq m)$ is

$$c_j = c_{min} + \Delta(j - 1).$$

Erev *et al.* defined $\Delta = 2(c_{\max}) / (m - 1)$ and set the two strategy-set parameters to $m = 101$ and $c_{\max} = 5$ (note that the symmetry assumption implies that $c_{\min} = -c_{\max}$).

Initial propensities.

A2^q. At time $t = 1$ (before any experience) the DM has an initial propensity to choose the jth cutoff.

Two initial propensities parameters were set, $S(1)$ and σ_i. To set these parameters Erev *et al.* defined $q_k(l) = p_k(1)S(l)$ where $S(1) = \sum_{j=1}^{m} q_j(t)$ and $p_k(1)$ is the probability that cutoff k will be chosen in the first round. They assumed that $S(1) = 3^*$ (the average absolute profit in the game, given randomly selected cutoffs) and determined $p_k(1)$ by the area (above cutoff k) under a normal distribution with a mean at the center of the two distributions and standard deviation $\sigma_i = 1.5$.

Reinforcement, generalization, and forgetting. The learning process is the result of the updating of the propensities through reinforcement, generalization, and forgetting.

A3^q. If cutoff k was chosen by the DM at time t and the received payoff was v then the propensity to set cutoff j is updated by setting

$$q_j\,(t + 1) = \max[\nu,(1 - \varphi)q_j(t) + G_k(j, R_t, (v(t)))],$$

where ν is a technical parameter that ensures that all propensities are positive, φ is a forgetting parameter $G_k(.,.)$ is a generalization function and $R(.)$ is a reinforcement function.

Erev *et al.* set $\nu = .0001$, $\varphi = .001$. The reinforcement function was set to $R_t\,(v(t)) = v(t) - \rho(t)$, where ρ is a reference point that is determined by the following contingent weighted average adjustment rule

$$\rho(t + 1) = \begin{cases} \rho(t)(1 - w^+) + v(t)(w^+) \text{ if } v(t) \geq \rho(t) \\ \rho(t)(1 - w^-) + v(t)(w^-) \text{ if } v(t) < \rho(t), \end{cases}$$

where $w^+ = .01$ and $w^- = .02$ are the weights by which positive and negative reinforcements affect the reference point.

The generalization function was set to:

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

where $F\,\{.\}$ is a cumulative normal distribution with mean c_k and standard deviation σ_g. Erev *et al.* set $\sigma_g = .25$.

The Relative Propensities Sum. The final assumption states the choice rule.

A4[q]. The probability that the observer sets strategy k at time t is determined by the relative propensities sum:

$$P_k(t) = q_k\,(t) \,/\, \left[\sum_{j=1}^{m} q_j\,(t)\right].$$

REFERENCES

Busemeyer, J. R, and 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.

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review,* **105**(2), 280–298.

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

Erev, I. and Roth A. E. (in press). Predicting how people play games: Reinforcement learning processes on games with unique mixed strategy equilibrium. *American Economic Review.*

Gopher, D., and Barzilai, O. (1993). The effect of knowledge levels on operators response to malfunctions and technical problems in the system. In *Proc. IEEE: Systems, Man and Cybernetics, October, France,* IEEE, New York.

Gopher, D., Itkin, T., Erev, I., Meyer, J., and Armony, L. (1995). *Top down and bottom up processes in perceptual games.* Presented at SPUDM-15, Jerusalem, Israel, August 1995.

Green, D. A., and Swets, J. A. (1966/1988). *Signal detection theory and psychophysics.* New York: Wiley (reprinted in 1988; Los Altos, CA: Peninsula).

Heinrich, H. W. (1931). *Industrial accident prevention: A scientific approach.* New York: McGraw-Hill.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica,* **47,** 263–291.

Kubovy, M., and Healy, A. F. (1977). The decision rule in probabilistic categorization: What it is and how it is learned. *Journal of Experimental Psychology: General,* **106,** 427–446.

Lee, W. (1963). Choosing among confusably distributed stimuli with specified likelihood ratios. *Perceptual and Motor Skills,* **16,** 445–567.

McAfee, R. B., and Winn, A. R. (1989). The use of incentives/feedback to enhance work place safety: A critique of the literature. *Journal of Safety Research,* **20,** 7–19.

McKenna, F. P.(1988). What role should the concept of risk play in theories of accidents involvement? *Ergonomics,* **31,** 469–484.

Moray, N. (1990). Designing for transportation safety in light of perception attention and mental models. *Ergonomics,* **33,** 1201–1213.

Petersen, D. (1988). *Techniques of safety management: A scientific approach.* New York: Allorney.

Reason, J. (1995). A systems approach to organizational error. *Ergonomics,* **38,** 1708–1721.

Roth, A., and Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior,* **3,** 3–24.

Roth, A. E., and Malouf, M. W. K. (1979). Game-theoretic models and the role of information in bargaining. *Psychological Review,* **86,** 574–594.

Selten., R. (1996). *Learning direction theory.* Presented at the 29th Annual Meeting of the Society for Mathematical Psychology, Chapel Hill, North Carolina, August 1996.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appelton Century Crofts.

Sperling, G., and Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. Boff, L. Kaufman, & J. Thomas (Eds.) *Handbook of perception and human performance. Vol. 1. Cognitive processes and performance.* New York: Wiley.

Starbuch, W. H., and Milliken, F. J.(1988). Challenger: Fine-tuning the odds until something breaks. *Journal of Management Studies,* **25,** 319–339.

Summala, H. (1988). Risk control is not risk adjustment: The zero-risk theory and its implications. *Ergonomics,* **31,** 491–506.

Swets, J. A., and Green, D. M. (1978) Applications of signal detection theory. In H. L. Pick, Jr., H. W. Leibowitz, J. E. Singer, A. Steinschneider, & H. W. Stevenson (Eds.) *Psychology: From research to practice* (pp.311–331). New York: Plenum Press.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs,* **2.**

Wilde, G. J. S. (1988). Risk homeostasis theory and traffic accidents: Propositions, deductions and discussion of dissension in recent reactions. *Ergonomics,* **31,** 441–448.

Received: June 11, 1997