Journal of Experimental Psychology
Vol. 57, No. 1, 1959

DECISION-MAKING BEHAVIOR IN A TWO-CHOICE UNCERTAIN OUTCOME SITUATION [1]

SIDNEY SIEGEL AND DONALD AARON GOLDSTEIN

Pennsylvania State University

The predictions people make when placed in a two-choice uncertain outcome situation have received considerable attention in recent years. In the classical situation, as first used by Humphreys (1939), *S* is asked to predict which of two events (e.g., illumination or nonillumination of a light) will occur after a signal stimulus. The two events occur with fixed but unequal probabilities, say π_1 and π_2, in a random sequence for a number of trials, and the occurrence of either is not contingent on *S*'s behavior. The *S* is instructed to do his best to predict which will occur. He is allowed to witness the event, and thus to determine for himself the correctness or incorrectness of his prediction.

There are two theoretical models of interest here which yield predictions about *S*'s behavior in a situation of this sort.

One is Estes' (1950, 1954; Estes & Burke, 1953) statistical learning model which yields the prediction that *S*s will learn to match their response ratios (the relative frequencies with which they predict each of the events) to the actual probabilities of occurrence of the events. The same prediction is yielded by the Bush-Mosteller (1955) model when certain restrictions are applied to the parameters of that model. A number of studies report findings in support of this prediction (Estes & Straughan, 1954; Grant, Hake, & Hornseth, 1951; Hake & Hyman, 1953; Jarvik, 1951).

The second model, formalized by Von Neumann and Morgenstern (1947), is a game-theoretic model. According to some interpreters, one prediction consistent with this model is that a person will learn to maximize the expected frequency of correct predictions. This is accomplished by his predicting the more frequent event on all trials. Evidence giving some support to this prediction may be found elsewhere (Galanter & Gerstenhaber, 1956; Goodnow, 1955; Toda, 1956).

From decision-making theory (Davidson, Suppes, & Siegel, 1957; Edwards, 1954), an hypothesis of *maximization of expected utility* may be drawn which will account for both sorts of prediction behaviors (matching response ratios and maximizing the expected frequency of correct predictions). The "utility" of an outcome is its subjective value. According to the decision-making approach, *S*'s choice behavior (here, his predictions) depends on certain conditions related to the reinforcement inherent in the situation. The general hypothesis is that a person will maximize expected utility in any case. It must be pointed out, however, that the components of the total utility vary in magnitude from one reinforcement situation to another, and thus the strategy to maximize expected utility varies.

It is reasonable to suppose that when *S* is in a situation in which the

[1] This investigation was supported by a research grant (M-1328) from the National Institute of Mental Health, U. S. Public Health Service.

only payoff attached to the outcomes is the satisfaction of having his prediction confirmed by the event or the dissatisfaction of having his prediction dis-confirmed by it, making a correct prediction of the rarer event has greater utility for *S* than making a correct prediction of the more frequent event. The *S* may be supposed to derive satisfaction from "playing a game with the machine," trying to "outwit" it. Moreover, it may be monotonous to make the same prediction on trial after trial over a long series, and predicting the rarer event may therefore satisfy *S*'s utility of relieving boredom. If it is true that among the components of *S*'s total utility in this "no-payoff" situation are the utility of correctly predicting the more frequent event, the (possibly greater) utility of correctly predicting the rarer event, and the utility of relieving boredom, then *S* may maximize his total satisfaction in this situation by choosing a "mixed strategy," and this may approximate matching his response ratio to the actual probabilities of occurrence of the two events. If such an account is correct, then a decision-making model and the Estes model could yield the same predictions concerning behavior in a two-choice no-payoff situation.

If, however, the utility attached to correct predictions is increased by a change in the conditions of the game (i.e., if there is potential satisfaction beyond that of having an event confirm one's prediction), or if the cost attached to an incorrect prediction is increased, then the prediction yielded by the hypothesis of maximization of expected utility diverges from the prediction yielded by the Estes model. That is, if the decision-making approach proposed here is correct, the introduction of systematic variation in the utilities attached to correct or incorrect predictions should be followed by systematic differences in *S*'s responses. As the utility of a correct prediction is increased or the negative utility of an incorrect prediction is increased, *S*'s prediction of the more frequent event should tend to 100%—i.e., he should tend to choose a "pure strategy" rather than a mixed one.

In the experiment designed to test this contention, three different experimental conditions (levels of reinforcement) were selected to effect systematic variation in the utility attached to correct and incorrect predictions. Under the "no-payoff" condition, the correctness or incorrectness of *S*'s predictions did not affect his monetary holding. Under the "reward" condition, the utility attached to a correct prediction was greater, for there was a monetary payoff for each correct prediction. Under the "risk" condition, the utility attached to a correct prediction was still greater, for there was a monetary loss for each incorrect prediction as well as a monetary gain for each correct prediction, and thus a correct prediction rather than an incorrect prediction had twice as great an effect on *S*'s monetary holding under the "risk" condition as under the "reward" condition.

The purpose of the procedure to be described was to provide data to test the hypothesis that *the asymptotic probability of* S*'s predicting the occurrence of the more frequent event in a two-choice uncertain outcome situation is a function of the level of reinforcement present in the situation, such that the probability of predicting the more frequent event will tend toward unity as the rewards (positive utility) and costs (negative utility) of correct and incorrect predictions are increased.*

That is, using Estes' notation (cf.

Estes, 1950), where $\bar{p}_1(\infty)$ is the mean asymptotic probability of predicting E_1 for a group of like *S*s, the hypothesis is that $\bar{p}_1(\infty)$ under risk $> \bar{p}_1(\infty)$ under reward $> \bar{p}_1(\infty)$ under no payoff.

Method

Subjects.—The *S*s were 36 male undergraduate students at Pennsylvania State University.

Apparatus.—Each *S* was run individually. For the experimental session, *S* was seated before a board to which three electric light bulbs were attached. Two, the "event" bulbs, were 60-w. bulbs set 8 in. apart. Directly between these and 6 in. below was a red 7 ½-w. bulb which served as the signal stimulus (warning light). On the table in front of *S*, between him and the board, were two push buttons set 4 in. apart.

The *E* was seated behind the board. Unseen by *S*, his push buttons were connected to bulbs at *E*'s desk. Depending on which button *S* pressed, one of the two bulbs at *E*'s desk would light, enabling *E* to record *S*'s predictions for the trial. As will be discussed, the sequence of "event" lights (right or left) for the successive trials was determined in advance of a session, and *E* followed this sequence in throwing a switch from his desk causing one or the other "event" bulb to light. None of *E*'s switching or recording equipment was visible to *S*.

For each trial, the signal stimulus (red bulb) would light and remain on for 3 sec. During this time, *S* made his prediction by pressing one of the buttons on the table before him. When the red bulb went off, one of the "event" bulbs would light and would remain on for 2 sec. A 2-sec. period intervened between trials.

Procedure.—The 36 *S*s were randomly assigned to three groups of equal size, the group assignment determining the condition under which each *S* would be run. The three conditions were No Payoff, Reward, and Risk.

Under the No Payoff condition, the reinforcement for each prediction consisted simply in seeing the outcome, i.e., seeing whether the right or left bulb lighted, and thereby determining whether one's prediction was confirmed or dis-confirmed. The No Payoff condition has been the "classical" situation for studies of human behavior in two-choice situations.

Under the Reward condition, the reinforcement for each prediction consisted in seeing the outcome and receiving 5 cents for each correct prediction. The 5-cent reward was given at the conclusion of each trial in which *S*'s prediction was confirmed.

Under the Risk condition, the reinforcement for each prediction consisted in seeing the outcome, receiving 5 cents if the prediction was confirmed, and losing 5 cents if it was dis-confirmed. At the conclusion of every trial, *S* either received or forfeited 5 cents, depending on whether his prediction had been correct or incorrect.

Two series of event sequences were prepared. One contained 75 lefts and 25 rights; the other was identical except that the lefts and rights were reversed, so that it contained 25 lefts and 75 rights. To counterbalance for the effects of possible right-left preference, half of the *S*s under each condition were presented with each series.

In the 100-trial series, the order of occurrence of the two events was randomized, by use of a table of random numbers, with two modifications: adjustments were made so that (*a*) there was no run of more than six successive occurrences of the more frequent event, and (*b*) within every 20-trial block the 3:1 proportion was maintained.

Each *S* was seen at an individual testing session in which a 100-trial series was presented. Each such session required about 30-min.

At the start of the session, standard instructions were given to all *S*s: "Your task in this experiment is to try to predict correctly which light will come on following the onset of the signal light. Three seconds after the signal light comes on, either the left or the right light will come on. As it is very important that you make a choice for each trial, please make your prediction as soon as possible after the signal light comes on. In any case, you must make your prediction in less than 3 sec.

"On the platform before you there are two buttons. If your prediction is that the left light will come on, press the left button. If your prediction is that the right light will come on, press the right button.

"Do you have any questions as to what you are to do? Let's run through what could happen. I will give you an example of each event."

In addition to these standard instructions and examples given to all *S*s, some additional special instructions were given depending on the condition under which the *S* was being run. The No Payoff *S*s were told simply: "Now, do your best to predict correctly which light will come on for every trial." The Reward *S*s received these additional instructions: "Now, every time you predict correctly you will win 5 cents; every time you predict incorrectly you will not win anything. At the end of the session all the money you have won will be yours to keep. That is, for every time you predict correctly you will win 5 cents, and all the money in your possession at the end of the session will belong to you. Now, if you make no correct predictions you will have won no money by the end of the game, but if you play carefully you should come out winning at the end of the session."

The additional instructions given to the Risk *Ss* were these: "Now, every time you predict correctly you will win 5 cents; every time you predict incorrectly you will lose 5 cents. We will start you off with 75 cents, and you will keep all of your winnings at the end of the session. That is, at the end of the session all the money in your possession will belong to you. Now, if you do not play carefully, you are likely to lose your starting stake and therefore end the session with no money. However, if you play carefully you could end the session with a considerable amount of money which would be yours to keep."

The instructions to both the Reward and the Risk *Ss* concluded with this statement:

"Let me make it quite clear: all the money you win you can keep. There are no strings attached, and please understand you will be paid the exact amount you win. This money does not belong to me; it comes from a research grant, so I am not concerned with how much or how little you win. The amount you win will depend entirely on how you play. Is that clear? All right. Now do your best to predict correctly which light will come on for every trial."

Every *S* under every condition was given 75 cents at the start of each session.

Results and Discussion

For the test of the hypothesis, *Ss*' final 20 predictions in the course of each 100-trial series were observed. The score for any *S* was the proportion of times he predicted (whether correctly or not) the occurrence of the more frequent event during those final 20 trials.

Table 1 shows the distribution of these scores for all *Ss* for the final 20 trials of the 100-trial series. Inasmuch as the research hypothesis was an ordered alternative hypothesis, these data were tested by the Jonckheere (1954) test, a nonparametric *k*-sample test against ordered alternatives. By this test, it is possible to reject the null hypothesis of no differences in favor of the alternative hypothesis at the $p < .00003$ level.

TABLE 1

Number of *Ss* Predicting More Frequent Event at Various Proportions During Final 20 Trials of First 100-Trial Series

Proportion of Predictions of More Frequent Event	Condition		
	No Payoff ($N = 12$)	Reward ($N = 12$)	Risk ($N = 12$)
.60	2		
.65	3		
.70	3	3	
.75	2	4	
.80	2	3	
.85		1	3
.90		1	1
.95			5
1.00			3

That is, the data confirm the prediction that *Ss* under the Risk condition will predict the more frequent light oftener than *Ss* under the Reward condition, and these in turn will predict the more frequent light oftener than *Ss* under the No Payoff condition. An interesting feature of the distributions shown in Table 1 is that there is no overlap between the scores of the Risk and No Payoff *Ss*.

The mean proportion of times that the more frequent light was predicted during the final trials of the first 100-trial series under the No Payoff condition was .70. Under Reward, the mean was .77. Under Risk, it was .93.

Four *Ss* were selected at random from the 12 under each condition and were run under the same condition for an additional 200 trials at a second session a week later. Although the number of cases is small, it is of interest to observe that these *Ss*' scores, during the final 20 trials of the second and third 100-trial series separately, each confirm the hypothesis, by the Jonckheere test, at well beyond the .01 level. For descriptive purposes, the mean proportions for these subgroups are presented in Table 2. The body of Table 2 gives the average proportion of trials on which the more frequent light was predicted by these *S*s in the various series under the various conditions. It is of some interest to observe that under both the No Payoff and the Reward conditions, these proportions increase from series to series (reaching $\pi_1 = .75$ for the No Payoff condition at the end of the third series), whereas under the Risk condition *S*s exhibited stable behavior, on the average, throughout the three series. It would seem that *S*s learn most rapidly under a condition of Risk (increased reinforcement).

The utility hinging on correct or incorrect predictions differed under the No Payoff, Reward, and Risk conditions, and thus the attempt to maximize expected utility led to different strategies under these conditions. Thus any attempt to construct a unified theory to account for the choice behavior to which both the Estes model and the game theoretic model have reference would have to consider, in addition to the utility of a correct prediction, other components of the total utility such as the specific utility of gambling, the negative utility of "boredom" (i.e., the result of always pressing the same button in predicting the same event over hundreds of trials), etc. If these components of the total utility are conceived as varying in magnitude under the experimental conditions, then the data which have been reported support the general hypothesis that *S*s' choice behavior reflects an attempt to maximize expected utility.

TABLE 2

Mean Proportion of Times the More Frequent Event was Predicted by a Subgroup of Four *S*s, Randomly Selected From Each Payoff Group, During Final 20 Trials of Each 100-Trial Series

Series	No Payoff	Reward	Risk
1–100	.69	.78	.95
101–200	.74	.85	.95
201–300	.75	.86	.95

Some investigators in this field, especially those influenced by game theoretic principles, have asserted that people who match their response ratios to the probabilities are acting "irrationally," because such people could increase their expected proportion of correct predictions by predicting the more frequent event on every trial. However, as Simon (1956, p. 271) has reminded us, one must bear in mind the distinction between *subjective* rationality (i.e., behavior that is rational, given the perceptual and evaluational premises of *S*) and *objective* rationality (rationality as viewed by *E*). The empirical fact that *S*s' proportions of choices of the more frequent event increase as reinforcement and risk are increased seems to show that, in their choice behavior, *S*s act as though they were attempting to maximize expected utility. If this is the case we can designate the mixed strategy behavior, in the No Payoff condition, as rational in that it will lead to the greatest expected payoff in total satisfaction.

Summary

The purpose of the experiment which has been reported was to test the hypothesis that the asymptotic probability of *S*'s predicting the occurrence of the more frequent event in a two-choice uncertain outcome situation is a function of the level of reinforcement present in the situation, such that the probability of predicting the more frequent event will tend toward unity as the rewards (positive utility) and costs (negative utility) of correct and incorrect predictions are increased. This hypothesis is drawn from the theory of decision making.

The *S*s' predictions were observed in a two-choice uncertain outcome situation under three conditions: No Payoff, Reward, and Risk. The proportion of times *S*s predicted the more frequent outcome differed under these three conditions in the predicted direction.

References

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Davidson, D., Suppes, P., & Siegel, S. *Decision-making: An experimental approach.* Stanford, Calif.: Stanford Univer. Press, 1957.

Edwards, W. The theory of decision-making. *Psychol. Bull.*, 1954, **51**, 380–417.

Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

Estes, W. K. Individual behavior in uncertain outcome situations: An interpretation in terms of statistical association theory. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes.* New York: Wiley, 1954. Pp. 127–137.

Estes, W. K., & Burke, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

Galanter, E., & Gerstenhaber, M. On thought: The extrinsic theory. *Psychol. Rev.*, 1956, **63**, 218–228.

Goodnow, J. J. Determinants of choice distributions in two-choice probability situations. *Amer. J. Psychol.*, 1955, **68**, 106–116.

Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, **42**, 1–5.

Hake, H. W., & Hyman, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45**, 64–74.

Humphreys, L. G. Acquisition and extinction of verbal expectation in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.

Jarvik, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, **41**, 291–297.

Jonckheere, A. R. A distribution-free *k*-sample test against ordered alternatives. *Biometrika*, 1954, **41**, 133–145.

Simon, H. A. A comparison of game theory and learning theory. *Psychometrika*, 1956, **21**, 267–272.

Toda, M. Guessing sequences under various conditions of payoff. *Jap. psychol. Research*, 1956, **4**, 11–23.

Von Neumann, J., & Morgenstern, O. *Theory of games and economic behavior.* (2nd Ed.) Princeton, N. J.: Princeton Univer. Press, 1947.

(Received February 17, 1958)