Journal of Experimental Psychology
1962, Vol. 63, No. 1, 42–49

PREFERENCE PROBABILITY BETWEEN GAMBLES AS A STEP FUNCTION OF EVENT PROBABILITY[1]

R. DUNCAN LUCE AND ELIZABETH F. SHIPLEY
University of Pennsylvania

This experiment is another of the studies arising ultimately from the utility ideas of Ramsey (1931) and von Neumann and Morgenstern (1947). It is based on a probabilistic analysis of choices among uncertain outcomes that rests upon two major assumptions. (Luce, 1959, Ch. 3). One, a general choice axiom, relates the choice probabilities for any set of three alternatives to those for its two-element subsets. Specifically, it postulates that if one of the three alternatives is eliminated, its probability is distributed between the other two in proportion to their probabilities in the original set of three. The second, a decomposition axiom (Luce, 1958; Luce & Raiffa, 1957) which has meaning only for uncertain outcomes, assumes that when S is choosing between certain pairs of gambles his preferences between the pure outcomes are statistically independent of his likelihood judgments between the chance events. More exactly, it says that if one knows the objective probability that one pure outcome is preferred to another and the objective probability that one event is deemed more likely to occur than another, then the probability of choosing between the gambles generated by these outcomes and events is given by the straightforward probability calculations under the assumption of independence. Although nothing is assumed in this postulate about the existence of subjective probability or utility functions as in classical utility theory, it is of the same spirit as the expected utility hypothesis, but much weaker. The precise mathematical statements of these two assumptions can be found in Luce (1959); they will not be repeated here.

[1] This research was supported in part by Grants NSF G-4506 and G-5544 from the National Science Foundation to Harvard University. The experimental work was carried out at the Laboratory of Social Relations, Harvard University.

Some psychophysical evidence supports the choice axiom (Anderson, 1959; Clarke, 1957); however, as yet, no direct preference tests have been made. Coombs (1958, 1959) presents indirect negative evidence, but there is some doubt about its meaning (Luce, 1959, p. 19). No data have yet been published on the decomposition axiom. On both counts, therefore, data are needed. Because these assumptions are difficult to test separately, and because together they imply a surprising result, we shall test only their conjunction.

Consider the one-person game:

$$\begin{array}{ccc} & \text{Option I} & \text{Option II} \\ \begin{matrix} p \\ \bar{p} \end{matrix} & \left[\begin{matrix} a \\ b \end{matrix}\right. & \left.\begin{matrix} c \\ d \end{matrix}\right] \end{array}$$

where a, b, c, and d are sums (possibly negative) of money such that $a > c > d > b$ and where p and $\bar{p}$ mean the occurrence and nonoccurrence, respectively, of a certain chance event. The S selects one of the two options and then, following his choice, the chance event is given an opportunity to occur. These two decisions—one by S, one by chance—determine a cell in the matrix, which contains S's payoff. Thus, if S chooses Option II and the chance event fails to occur (row $\bar{p}$), the payoff to S is d.

Now, holding the payoffs fixed, consider S's choices as we vary the event probability. If the event is certain to occur—or so S believes—then it is only sensible for him always to choose Option I because chance will choose Row 1 and, by assumption, S prefers a to c. On the other hand, if the event is certain not to occur, S surely will never choose Option I because $b < d$. And as the event probability is varied from zero to one, S's probability of choosing Option I must change from zero to one. The question is: what is the nature of this transition?

Classical economic thought has assumed a simple discontinuity, but most psychologists doubt that this is correct, and the evidence (Chipman, 1960; Coombs, 1958, 1959; Davidson & Marschak, 1959; Davidson, Suppes, & Siegel, 1957; Mosteller & Nogee, 1951; Papandreou, 1957), although open to other interpretations, is generally viewed as being negative. Assuming that the function is not a simple discontinuity, the most plausible a priori conjecture is a smooth ogival function, so familiar from psychophysics. Plausible, yes, but not possible for all choices if the two assumptions mentioned above are correct. For payoff matrices where one row dominates the other (e.g., $a>b$ and $c > d$), they imply that the choice probability varies from zero to one by a series of jumps. That is to say, the function must be a step function, of which the simple discontinuity is a special case.[2] When neither row dominates the other, i.e., when the two largest payoffs lie on a diagonal, no theoretical prediction has been derived. It is not known whether steps are again necessary or whether continuous functions are possible.

[2] It should be noted that this conclusion, which is "global" in nature, follows from repeated application of Luce's (1959, p. 86) Theorem 13, which is "local" in nature. The theorem requires that various probabilities of choice be different from 0 and 1, which is what gives it its local character, but this assumption need not hold for any two arbitrarily chosen gambles of an experiment designed to test the step-function hypothesis. Thus, there can be no question of looking to the data to see whether prerequisites of the step-function hypothesis are satisfied.

Although this prediction of steps is qualitatively startling and sharply at variance with most psychologists' intuitions, it is uncertain whether it can be successfully tested—not only would negative results be completely ambiguous, but seemingly positive ones are open to other interpretations. Basically, there are three difficulties. First, the theory states neither how many steps to expect nor where to find them. If there are enough, say 8 or 10, reasonably spaced over a sufficiently narrow region of event probabilities, then a vast amount of data is needed to tell a step function from a smooth ogive. Second, to estimate the choice probabilities, the choices must be presented repeatedly, allowing all the well-known dangers of systematic changes over time. And third, even if by good luck and design these troubles are avoided, one must still live with the inevitable and considerable binomial variance of our observations which, on the one hand, may tend to mask the steps of a step function (unless they are broad) and, on the other hand, may introduce at least small steps into data arising from an ogive.

Severe though these difficulties seem to be, it seemed worthwhile to conduct the obvious experiment. The payoff, should the data hint at positive results, seemed sufficiently large to offset both the slim chance that the steps are there and the inherent uncertainties of knowing it even if they are.

Method

Subjects.—Without specifying the exact nature of the work, five Harvard College students were solicited to participate in an experiment for which they would be paid both an hourly wage and a bonus. After he had indicated interest in the work, each was asked if he had qualms about gambling; none did.

Apparatus.—The 2×2 payoff matrices were printed on small white cards. Each column defined a gamble, and so an option for S; each row corresponded to one of two complementary events. Packs of these cards, face down, were mounted in a rack before S. Turning a card over to expose the matrix tripped a microswitch which in turn activated a pen of an Esterline Angus recorder. Two buttons, labeled to correspond to S's two options, were located in front of each S and were wired to the pen recorder.

The events were specified in terms of the faces shown by five dice. The possible outcomes were ranked into hands, much as in poker except that flushes could not occur, there being no suits, and straights were defined to be of the lowest rank. A hand known as the house's hand was printed on each card. One event was defined to have occurred when the dice exhibited a hand, known as S's hand, higher than the house's; the complementary event occurred when S's hand was the same as or less than the house's. Each S was provided with an ordered list of the 252 hands that might arise and of the probabilities of beating each of them. The events were produced by spinning a wire cage that contained the five dice.

Procedure.—Prior to the experiment, E explained both the procedure and the bonus system. The payoffs in the matrices were treated as points to be won or lost on each throw of the dice. At the end of the experiment the accumulated points would then be used to determine each S's share of a $200 bonus. They were told that a point was roughly equal to 1 cent, which was true.

An experimental trial was as follows. When E gave a ready signal to begin the trial, each S turned over the next card in his pile, examined the alternatives, and indicated his choice by pressing one of his two buttons. The time when the card was turned over, his choice, and the time it was made were all recorded. After all Ss had made their choices, one of them—the same one throughout an experimental session—spun the cage and called out the resulting hand. Each S then recorded on a sheet provided the points won or lost on the trial. They knew that on any given trial the payoff matrix and the house's hand differed among them, but that throughout the experiment they would all be presented with the same six matrices. They were also led to believe, although this was not true, that they would all see the same house hands.

Four preliminary sessions were run to determine for each S-matrix combination the range of events in which S's choices were uncertain. The experiment itself consisted of daily sessions of 150 trials, lasting about 1 hr., for 33 days. Because some of the initial choices of event ranges for S_1 were inappropriate, he was run three extra sessions. For the first few weeks, sessions occurred only when all 5 Ss could attend; however, S_1 became ill, which made it necessary to relax this requirement. It seemed that the larger the group, the more enjoyable the session; some sessions at the end with a single S were considered dull. On the whole, however, if noisy rivalry is any sign, the Ss were interested and involved.

TABLE 1

PAYOFF MATRICES STUDIED

1		2		3	
I	II	I	II	I	II
+31	+22	+14	+5	−4	−13
+10	+16	−7	−1	−25	−19
4		**5**		**6**	
I	II	I	II	I	II
+33	+17	+11	−5	−5	−21
+12	+23	−10	+1	−26	−15

Conditions.—The six payoff matrices used are shown in Table 1. The rational decision rule is the same for all: choose Option I if the probability of the first row event exceeds .4, and choose Option II if this probability is less than .4. The first three matrices differ from each other only by a positive or negative constant added to all entries. The entries are all positive in the first matrix, mixed in the second, and negative in the third. The last three matrices are similarly related. The first three meet the necessary inequalities so that a step function is predicted in response to them. For the second set of three, no prediction from the theory is known. They were run to see whether Ss appear to react differently to the two classes.

During the 4-day preliminary phase, seven events with probabilities from .2 to .8, spaced in .1 steps, were used. Each of the 42 event-matrix combinations was presented from 9 to 13 times. These data were used to estimate the approximate event probability for which S's choice probability was .5 for each matrix. A range from .1 below to .1 above this point was explored in the experiment. As mentioned above, some mistakes about S_1's ranges were made, but they were rectified shortly after the experiment began. Aside from that, all save one of the other ranges were appropriate.

Fifteen hands were selected in these .2 probability ranges so that the steps between adjacent ones varied from .020 probability at the edges to .008 in the center of the range. Because we did not want *S*s to be aware of differential treatment, events spaced at .05 intervals from .25 to .75 were also presented to all *S*s. Nonetheless, S_4 and S_5 realized that the other *S*s were getting different house hands, and they questioned *E* about this. They were assured that we would distribute the bonus on the basis of the number of points earned relative to the maximum possible.

Each of the 90 event-matrix combinations appeared once every 90 trials for a total of 50 presentations during the experiment. Within a block of 90, their order was randomized. The 36 event-matrix combinations that were added to this basic design to convince *S*s that they were being treated alike were each presented five times. Rows or columns were interchanged randomly on successive presentations of the same event-matrix pair.

Results and Discussion

Plots of the estimated choice probabilities vs. the theoretical event probabilities for the six payoff matrices of the S_1, S_2, and S_3 are shown in Fig. 1. To save space, the corresponding plots for S_4 and S_5 are not presented because in all cases they are almost perfect simple discontinuities. For S_4 the break occurs at event $p = .4$, which is the rational decision, and for S_5 it occurs at $p = .5$. Thus, the data from S_4 and S_5 are consistent both with the step function hypothesis and with the classical economic assumption of a simple

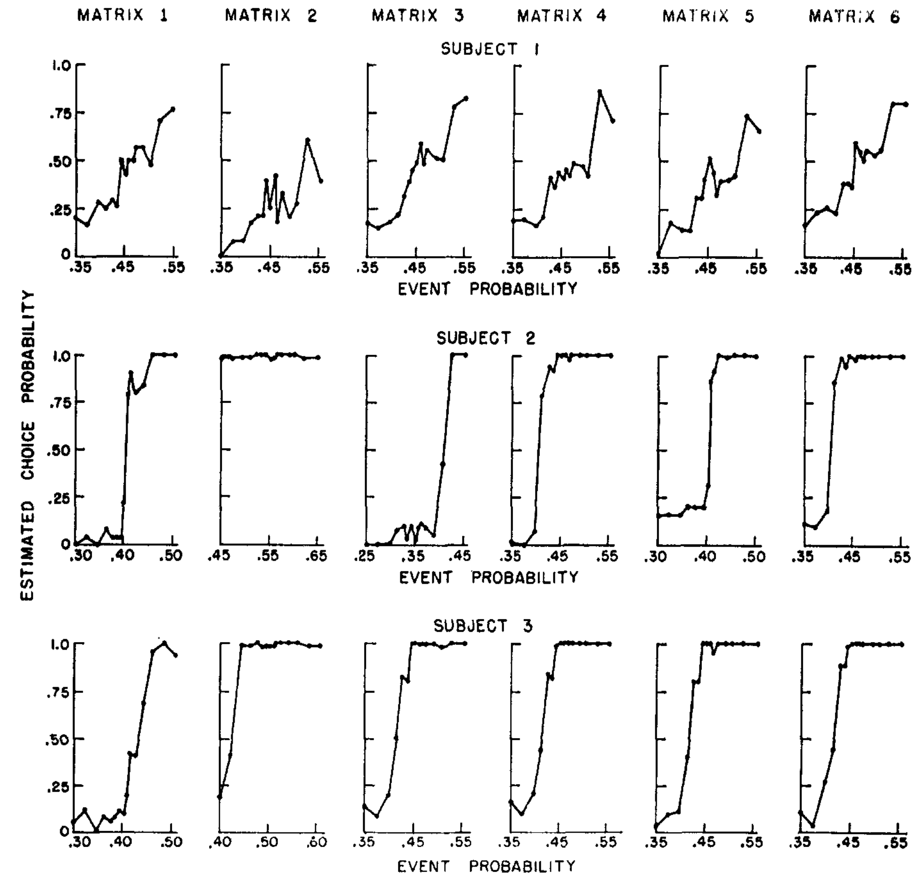

FIG. 1. Estimated choice probabilities vs. event probabilities for S_1, S_2, and S_3 on the six different payoff matrices. (Each point is based upon 50 observations.)

discontinuity, but they are not consistent with the ogival hypothesis (except, of course, in its limiting form). The data for *S*s 1–3 are certainly inconsistent with the assumption of a simple discontinuity, but whether steps or ogives underlie these data is less clear.

Because we do not know how to reduce this question to a well formulated and solved statistical problem, the reader ultimately will have to make his own decision. The following considerations may be helpful.

It is clear that the data exhibit plateaus, some of which are broad (particularly in S_1, Matrices 1–6; S_2, Matrices 3 and 5; and S_3, Matrix 1). There is also no question that the edges of these flats are rounded, but considering all the possible sources of blurring, this is the least to expect if the step function hypothesis is correct. So the only problem we need face is whether these plateaus could have arisen from a continuous, ogival function. We have thought of only two possible ways: binomial variability and an abrupt change in the location of the ogive. Both are discussed.

Three arguments can be brought against the binomial variability explanation. The first is simply the subjective impression that each *S* exhibits a characteristic style of behavior over the six matrices. For example, in S_3 there is a small jog at about $p = .425$ in five of the six matrices. The style seems to persist even when different ranges of event probabilities were employed (see S_2 and S_3). If the flat regions really result only from sampling variability, no internal *S* consistency, such as there seems to be, should exist.

Second, if variability accounts for the plateaus, then they should not stand up very well when the data are divided into two parts. Separating the even and odd observations for each point, we obtain plots such as those in Fig. 2.[3]

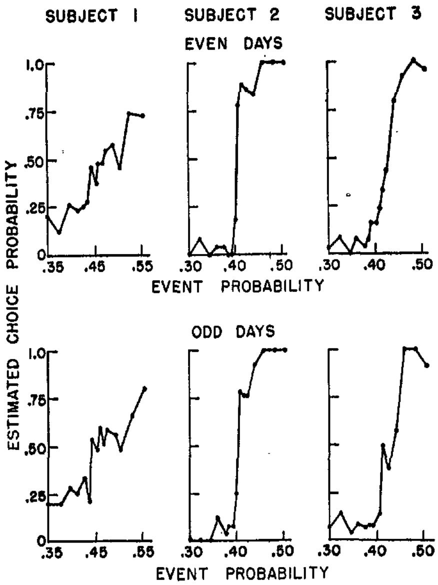

FIG. 2. Estimated choice probabilities vs. event probabilities for S_1, S_2, and S_3 on Payoff Matrix 1 for the even and the odd days. (Each point is based upon approximately 25 observations.)

Note that at least one of the narrower plateaus vanishes, as might be expected, but the broader ones remain.

Third, in the absence of more sophisticated statistical procedures, we can generate Monte Carlo data to see whether, by chance, plausible ogives could be expected to generate reversals of the magnitude seen in the data. To this end, we had the following problem programed for the Univac computer at the University of Pennsylvania.[4] Given $0 < p_1 < p_2 < \cdots < p_{15} < 1$, use random numbers generated in the machine to obtain N (in our case, 50) independent binomial observations for each p_i. Let N_i be the number of successes. For each lag k, $k = 1, 2, \ldots, 8$,

[3] As it is impractical to present all 36 of these curves, and as there seems to be no particularly rational way to choose a few from among them, we simply have presented the first matrix for each *S*. The same choice is made in Fig. 3.

[4] We wish to thank Donald Norman and A. Martin Wall for preparing the program and carrying out the runs and we would also like to express our appreciation to the Computer Center of the University of Pennsylvania for its cooperation.

TABLE 2

Properties of the Logistic Functions $P=1/1+e^{a\sigma})$ Used to Generate the Monte Carlo Data

Ogive Number	σ for which $P = 1/2$	jnd	σ Range
1	.45	.06	.35–.55
2	.40	.02	.30–.50
3	.40	.02	.35–.55

store the number of differences $N_{i+k}-N_i$ which are zero or negative. Repeat this M times (we used $M = 200$) and read out, for each k, how many cases exhibited 0, 1, . . ., and 15-k reversals, i.e., zero and negative differences.

Three sets of probabilities p_i were selected as follows. Three logistic curves, $1/(1 + e^{a\sigma})$, were chosen to approximate most of the data curves; they have the features shown in Table 2. By letting σ run over the event probabilities used in the experiment, a set of 15 probabilities was generated from each logistic.

For each S and each matrix, we determined for each lag k the number of times the choice proportion was the same or smaller for an event k steps above another event. Then, using the Monte Carlo data for the best fitting of the three ogives, we found the percentage of times that many or more reversals of lag k occurred in the Monte Carlo runs. Thus, for example, S_2 in Matrix 4 exhibited six reversals of lag 2, and the Monte Carlo runs for Ogive 3 showed five cases out of 200 with six or more reversals at lag 2. So the entry for this case is $5.100/200 = 2.5\%$. These estimated probabilities, which are summarized in Table 3, measure how likely it is that reversal structure of our data could have arisen by chance from the ogive. It should be kept in mind that the eight entries per row are not independent.

The patterns for S_2 and S_3 seem unambiguous; it is most unlikely that our data could have arisen from a logistic with a .02 jnd. The pattern is less clear for S_1, but we note that except for Matrix 3 there is always at least one lag for which the observed number of reversals has 5 chances or less in 100 of occurring if the ogive is the true function.

On the basis of these considerations, we doubt that the observed plateaus can have resulted simply from ogives plus sampling variability.

A second way a continuous ogive might give rise to a plateau is for the center of the ogive to be located at one event probability on some trials and at another on the remainder. If the absolute value

TABLE 3

Percentage of 200 Monte Carlo Runs in Which the Number of Reversals of Lag k Equaled or Exceeded the Number Observed

S	Matrix	Ogive	1	2	3	4	5	6	7	8
1	1	1	2	97.5	3.5	34.5	*	2	*	*
	2	1	77.5	40	3.5	0	0.5	0	*	0
	3	1	77.5	40	28	34.5	*	*	*	*
	4	1	5	76.5	28	34.5	18.5	2	*	0
	5	1	17.5	40	3.5	6	0.5	*	*	*
	6	1	5	14	28	34.5	18.5	*	*	*
2	1	2	0	0	0	0	0	*	*	*
	4	3	0	2.5	1	0.5	1.5	0	0	2
	5	2	0	0	0	0	*	*	*	*
	6	3	0	2.5	0.5	0.5	1.5	2.5	8	2
3	1	2	14.5	0 ·	0.5	0	0	0	0	*
	3	3	0	2.5	0.5	0.5	0	0	0	2
	4	3	0	2.5	0.5	0.5	1.5	2.5	8	*
	5	3	1.5	2.5	0.5	0.5	1.5	0	0	2
	6	3	0	2.5	0.5	0.5	1.5	2.5	8	*

Note.—The asterisk indicates that in these cases no reversals were observed, hence that number was equaled or exceeded in all Monte Carlo runs.

of the difference in locations is large relative to the breadth of the ogive, then the expected data would exhibit a single intermediate step with rounded edges. With three discrete locations for the ogive, two rounded steps can result. And so on. Such discontinuous shifts of location might, for example, correspond to major shifts in S's opinion about the rational breaking point. If these abrupt changes in location occurred irregularly from day to day during the experiment, we know no way of detecting them; however, if the shifts occurred only once or twice during the experiment, then, for example, the data from the first half of the experiment should differ appreciably from those from the second half. Such plots are shown in Fig. 3; the horizontal features of the plots are substantially the same in the two halves.

Although there is no evidence for the sort of shift that would produce plateaus, this figure does suggest that there may have been systematic changes in the underlying probabilities during the ex-

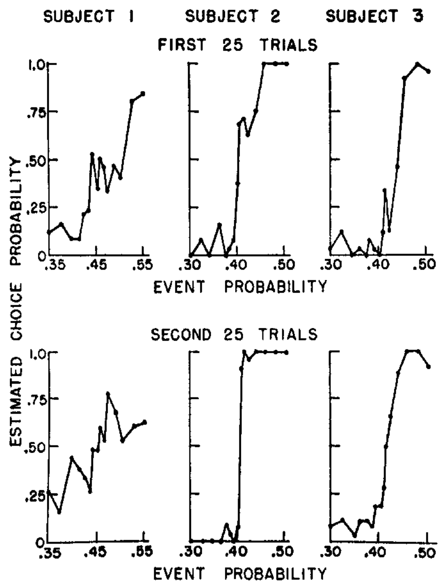

FIG. 3. Estimated choice probabilities vs. event probabilities for S_1, S_2, and S_3 on Payoff Matrix 1 for the first 25 and the second 25 trials. (Each point is based upon 25 observations.)

TABLE 4

P-VALUE COMPARISONS BETWEEN FIRST AND SECOND HALVES OF THE DATA

S	Matrix					
	1	2	3	4	5	6
1	>.05	>.05	.01–.02	>.05	.02	<.01
2	>.05	.01	<.01	>.05	.01	>.05
3	<.01	<.01	.02	.02–.05	.05	.02

Note.—P values are based on Wilcoxon matched-pairs signed rank test.

periment. To check this, the choice proportions for the first and second halves of the data were compared for each S-matrix combination by the Wilcoxon matched-pairs signed rank test. The P levels are summarized in Table 4, where we see that in over half the cases there are significant trends. The effects are not, however, very large. Their main consequence would be to increase the variance of the data slightly beyond binomial.

On the basis of the preceding arguments, we are inclined to view the data as supporting the experimental hypothesis that the choice probability for certain pairs of gambles is a step function of the event probability.

It will be recalled that in contrast to matrices, such as 1–3, in which one row dominates the other, no corresponding theoretical prediction about the response probabilities is known for matrices, such as 4–6, in which the two largest payoffs lie on a diagonal. This may just be our ignorance or it may be that the theory is simply noncommittal about these matrices. In any event, the data fail to suggest any behavioral differences. If the predicted steps actually exist, they also appear when neither row dominates the other; if not, then neither class of matrices exhibits steps.

Furthermore, there is no suggestion that the sign of the payoffs differentially affects the qualitative features of the behavior.

Finally, we turn to the other class of data recorded, response times. It is often held that response times, although ill understood, are closely related to response probabilities, and so it is conceivable that they too might exhibit discontinuities. The means and variances of the response times were plotted as functions both of the response probabilities and of the event probabilities for each *S*-matrix pair. These plots are not presented because, so far as we could see, there is no consistent picture. Much of the possible detail is masked by variability, and those trends that do show through differ both within *S*s and within matrices. Probably one cannot expect to make subtle uses of response times in experiments, such as this one, where no strong external time pressures are imposed.

Summary

In this experiment, each *S* chose repeatedly between pairs of gambles. His payoff on each trial was determined by his choice together with the outcome of a chance event which was completely independent of his choice. Our purpose was to test the mathematically derived prediction that, when certain inequalities are satisfied by the payoffs, *S*'s choice probability is a step function of the event probability. As an alternative hypothesis we selected the intuitively plausible one that the function relating these two probabilities is ogival. Although qualitatively these hypotheses contrast sharply, various a priori considerations make it difficult to discriminate between them.

Data were collected from 5 *S*s. Each was given six payoff matrices with 15 events per matrix and 50 observations per event. Two *S*s exhibited almost perfectly discontinuous functions, and so they were (trivially) consistent with the step function hypothesis and only consistent with the ogival hypothesis in its limiting form. The other 3 *S*s produced what can best be called lumpy functions. Arguments were adduced to suggest that neither variability coupled with an underlying ogive nor abrupt shifts in the location of an ogive are sufficient to explain the data. No other ways have been suggested whereby an ogive could generate data having plateaus.

The tentative conclusion is that the data support the step function hypothesis more strongly than the ogival one, and that indirect confirmation is given to the assumptions underlying the prediction.

References

Anderson, C. D. The constant-ratio rule as a predictor of confusions among visual stimuli of brief exposure duration. Technical Note, 1959, Hearing and Communication Laboratory, Indiana University.

Chipman, J. S. Stochastic choice and subjective probability. In Dorothy Willner (Ed.), *Decisions, values, and groups.* New York: Pergamon, 1960. Pp. 70–95.

Clarke, F. R. Constant-ratio rule for confusion matrices in speech communication. *J. Acoust. Soc. Amer.*, 1957, **29,** 715–720.

Coombs, C. H. On the use of inconsistency of preferences in psychological measurement. *J. exp. Psychol.*, 1958, **55,** 1–7.

Coombs, C. H. Inconsistency of preference as a measure of psychological distance. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories.* New York: Wiley, 1959. Pp. 221–232.

Davidson, D., & Marschak, J. Experimental tests of a stochastic decision theory. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories.* New York: Wiley, 1959. Pp. 233–269.

Davidson, D., Suppes, P., & Siegel, S. *Decision making.* Stanford: Stanford Univer. Press, 1957.

Luce, R. D. A probabilistic theory of utility. *Econometrica*, 1958, **26,** 193–224.

Luce, R. D. *Individual choice behavior: A theoretical analysis.* New York: Wiley, 1959.

Luce, R. D., & Raiffa, H. *Games and decisions.* New York: Wiley, 1957.

Mosteller, F., & Nogee, P. An experimental measurement of utility. *J. pol. Econ.*, 1951, **59,** 371–404.

Papandreou, A. G. A test of a stochastic theory of choice. *U. Calif. Publ. Econ.*, 1957, **16,** 1–18.

Ramsey, F. P. *The foundations of mathematics.* New York: Harcourt, Brace, 1931.

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

(Received October 17, 1960)