EXPERIMENTAL TESTS OF A STOCHASTIC DECISION THEORY

BY

DONALD DAVIDSON AND JACOB MARSCHAK

TECHNICAL REPORT NO. 17

JULY 25, 1958

PREPARED UNDER CONTRACT Nonr 225(17)

(NR 171-034)

FOR

OFFICE OF NAVAL RESEARCH

BEHAVIORAL SCIENCES DIVISION

APPLIED MATHEMATICS AND STATISTICS LABORATORY
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

EXPERIMENTAL TESTS OF A STOCHASTIC DECISION THEORY

By

Donald Davidson and Jacob Marschak*/

INTRODUCTION

Common experience suggests, and experiment confirms, that a person does not always make the same choice when faced with the same options, even when the circumstances of choice seem in all relevant respects to be the same. However, the bulk of economic theory neglects the existence of such inconsistencies; and the best known theories for decision making, for example, those of von Neumann and Morgenstern [12] or Savage [15], base the existence of a measurable utility upon a pattern of invariant two-place relations, sometimes called 'preference' and 'indifference.' This raises a difficulty for any attempt to use such theories to describe and predict actual behavior.

---

*/ This paper will appear as part of a symposium on measurement to be published by Wiley under the editorship of C. West Churchman.

Research undertaken by the Applied Mathematics and Statistics Laboratory, Stanford University, under contract Nonr 225(17), NR 171-034, with the Office of Naval Research, and by the Cowles Commission for Research in Economics under contract Nonr-358(01), NR 047-006 with the Office of Naval Research.

The authors were helped by discussions with G. Debreu, E. Fels, L. Hurwicz, R. Radner, H. Raiffa, R. Savage, R. Summers and P. Suppes.

A number of ways of meeting the difficulty may be mentioned: (1) It is possible to insist on the normative status of the theory and construe all deviations as evidence of error on the part of the subject. (2) One may defend the descriptive accuracy of the theory and argue that it has been incorrectly interpreted; for example, by wrongly identifying two options (say winning one dollar at time t and winning one dollar at time $t + 10$ minutes) as the same. (3) One may interpret every case of inconsistency as a case of indifference: if the subject has chosen a rather than b but soon afterwards chooses b rather than a this is interpreted as indifference between those two objects; if he chooses a rather than b, b rather than c, and c rather than a this is interpreted as indifference between those three objects. In empirical application this approach would probably make indifference all-pervasive. (4) An alternative approach is to define preference and indifference in terms of probabilities of choice. Mosteller and Nogee, in testing the von Neumann and Morgenstern axioms, considered a subject indifferent between two options when he chose each option half the time [11]; Edwards [5] has also used this method. In this approach probabilities of choice do not enter the formal axiomatic development. (5) A fifth strategy, explored in this paper, incorporates probabilities of choice into the axiomatic structure, and exploits their properties in scaling utilities.

I. PRIMITIVE AND DEFINED NOTIONS

We now introduce various concepts needed for the subsequent discussion. It should be emphasized that strictness has in many places in this paper been sacrificed to perspicuity; we trust that the knowing reader can make the corrections needed for formal accuracy. First we list the primitive notions:

1. A set A of alternatives.[*] A may include wagers (choices involving risk) as well as sure outcomes. In Sections I and II we shall treat alternatives quite generally. In Section III we shall use special properties of wagers.

2. The probability $P(a,b)$ that the subject, forced to choose between a and b, chooses a. We assume in what follows, for every a and b in A:

(a) $P(a,b) + P(b,a) = 1$,

(b) $P(a,b)$ lies in the open interval $(0,1)$.

In a fully formalized exposition these assumptions would appear as axioms or theorems; in the present paper we shall sometimes leave these assumptions tacit. Under a natural interpretation 2(a) has empirical content: it implies that when a subject is asked to choose between a or b, he always chooses a or b. Normally, we are not interested in testing 2(a); rather, we attempt to make it true by enforcing a choice. Therefore we may want to state our experimental hypothesis as follows: if 2(a) is true for a given subject, then the other axioms hold. If 2(a) fails for a subject we then reject the subject, not the hypothesis.

---

[*] We use the word 'alternative,' as is fairly common in the literature of decision theory, to mean one of two or more things or courses between which a person may choose.

Or we may want to include 2(a) in the hypothesis and reject the hypothesis for a subject who refuses an offered choice. For the experiments reported here the issue is academic. All subjects were docile.

For the case where $a = b$, 2(a) has the consequence $P(a,a) = 1/2$. Formal convenience dictates that we not exclude this case although we give it no empirical meaning.

Before commenting on 2(b) it will be useful to give some definitions.

Definition 3. a is absolutely preferred to b if and only if $P(a,b) = 1$. This concept corresponds to the psychologists' "perfect discrimination."

Definition 4. a is stochastically preferred to b if and only if $1/2 < P(a,b) < 1$.

Definition 5. a and b are stochastically indifferent if and only if $P(a,b) = 1/2$. Since we use the word 'indifferent' in no other sense, we often omit the word 'stochastically.'

Definition 6. c is a stochastic midpoint between a and b if and only if $P(a,c) = P(c,b)$.

In situations in which it is natural to apply the theory, it is obvious that cases of absolute preference occur, violating 2(b). In particular, one would expect that when a and b denote respectively, "receiving a dollars" and "receiving b dollars" (or, for that matter, m or n units of some commodity) then, $m > n$ implies $P(a,b) = 1$. More generally, if m_1, n_1 are amounts of some commodity and m_2, n_2 are amounts of a second commodity, and $m_1 > n_1$, $m_2 \geq n_2$, then the alternative consisting of receiving m_1 and m_2 will be absolutely preferred to the alternative n_1 and n_2. This extends also to bundles consisting of 3 or more commodities.

In the experimental testing of stochastic theories of choice various devices may be used to avoid comparisons of alternatives which yield absolute preferences. Papandreou et al [13], using appropriate commodity bundles, avoided cases of the sort just mentioned. The methods used for avoiding comparisons apt to generate absolute preference in the experiments reported here will be discussed presently.

So long as the assumption stated in 2(b) remains in force, it is not enough merely to avoid comparing alternatives one of which is absolutely preferred to another; the set A of alternatives to which the theory applies must contain no two such alternatives. While we have no solution on hand, we shall mention in the next section the possibility of modifying the formal system to eliminate dependence on assumption 2(b).

II. GENERAL STOCHASTIC THEORY OF CHOICE

An important aspect of a general stochastic theory of choice lies in the fact that without specifically considering wagers it is possible to obtain forms of measurement stronger than a mere ordering by imposing plausible conditions on probabilities of choice. When conditions of sufficient strength are satisfied it is possible to interpret a comparison of probabilities as a comparison of differences in subjective value or utility. This idea is captured in a general form by the following definition:

Definition 1. For a given subject, a real valued function u is called a utility function on A (in the sense of Definition 1) if and only if, for every a, b, c and d in A,

$$P(a,b) \geq P(c,d) \text{ if and only if } u(a) - u(b) \geq u(c) - u(d).$$

The technique of building a subjective scale on the basis of frequency of discriminated differences is common in psychophysics since Fechner [6]; however, the emphasis in psychophysics on relating the subjective (sensation) scale to a physical continuum (which is not assumed in utility measurement) tends to obscure the analogy. Discussion of the relation between psychophysical scaling and utility measurement will be found in Marschak [10] and Luce [8], [9].

There is a much used adage in psychophysics which may be taken as suggesting the principle underlying Definition 1 above: 'equally often noticed differences are equal [on the sensation scale] unless noticed always or never' (ascribed by Guilford [7] to Fullerton and Cattell). The final phrase of this adage enters a caveat which is clearly as pertinent in utility as in sensation measurement for, in our terms, the caveat concerns the case of absolute preference. Consider the case where $P(6¢, 5¢) = 1 = P(\$50000., \$0)$ and hence, by Definition 1, $u(6¢) - u(5¢) = u(\$5000.) - u(\$0.)$, which is intuitively absurd. The difficulty created by the existence of absolute preferences is thus clear. The approach to a solution which suggests itself is to add to Definition 1 the caveat 'provided neither $P(a,b)$ nor $P(c,d)$ is equal to 0 or 1.' This would require modification of the axiomatic conditions needed to prove the existence of a utility function. We have not attempted to carry out this modification, which may well not be trivial.

We now consider what conditions are sufficient for the existence of a utility function (in the sense of Definition 1). Fortunately in approaching this question we are able to depend on previous work due to the fact that any theory which makes essential use of a four place relation comparing intervals may, with fairly trivial modifications, be reconstrued as a theory in which the atomic sentences are all of the form '$P(a,b) \geq P(c,d)$' as demanded by Definition 1.*/

What constitutes sufficient conditions for the existence of a utility function depends, in part, on the nature of the set A. We therefore consider several cases.

(a) The set A contains a known finite number n of alternatives, $a_1, \ldots, a_n$. In this case it is always possible although perhaps tedious to stipulate conditions on the probabilities $P(a_i, a_j)$ necessary and sufficient for the existence of a utility function. A simple example (for $n = 3$) will be treated fully later. In general, it suffices, because of I.2(a), to consider those probabilities $P(a,b)$ that are $\geq 1/2$; a given complete ordering of these numbers yields, by Definition 1, a sequence of $n(n-1)/2$ inequalities of the form

$$u(a_g) - u(a_h) \geq u(a_i) - u(a_k) \geq \ldots \geq 0$$

---

*/ The modifications may allow for the special properties of probabilities, and for the fact that '$P(a,b) \geq P(c,d)$' compares signed intervals while the quaternary relations taken as primitives in some theories compare unsigned intervals.

involving a set of only n distinct unknowns viz., the utilities of the n alternatives. Whether these inequalities have a solution can be answered separately for each of the possible $[n(n-1)/2]!$ orderings of the probabilities.

(b) The set A contains an arbitrary number of alternatives which are equally spaced in utility (such that for every a, b, c and d in A, if a and b are adjacent in utility*/ and c and d are adjacent, then $P(a,b) = P(c,d)$). The axioms are an obvious modification of the axioms on page 31 of Davidson, Suppes and Siegel [3].

(c) It will be convenient to give two definitions. The first we owe to Professor Patrick Suppes.

Definition 2. A set A of alternatives is stochastically continuous if and only if it meets the following three conditions for every a, b, c and d in A:

(i) there exists a stochastic midpoint between a and b;

(ii) if $P(c,d) > P(a,b) > 1/2$ then there exists a g such that $P(c,g) > 1/2$ and $P(g,d) \geq P(a,b)$;

(iii) (Archimedean condition) if $P(a,b) > 1/2$ then for every probability q such that $P(a,b) > q > 1/2$ there exists a positive integer n such that $q \geq P(a,c_1) = P(c_1,c_2) = \ldots = P(c_n,b) > 1/2$.

---

*/ Let $P(a,b) \geq 1/2$; then a and b are said to be adjacent in utility if $\overline{P(a,b) \leq P(a,c)}$ for every c with $P(a,c) \geq 1/2$.

Definition 3. The quadruple condition is satisfied if and only if for every a, b, c and d in A, if $P(a,b) \geq P(c,d)$ then $P(a,c) \geq P(b,d)$.

It follows immediately from Definition 1 that if a utility function exists on A then the quadruple condition is satisfied in A. However, we are now in a position to assert more:

Theorem 4. If A is stochastically continuous then a utility function exists if and only if the quadruple condition is satisfied.

A proof of this theorem will not be given here. The general line of demonstration is as follows: Suppes and Winet have given an axiomatization of utility based on a primitive concept which compares utility differences and have proven that if certain axioms on a relation between two pairs of alternatives hold then utility-differences can be defined, and hence, a function analogous to a utility function (in the sense of Definition 1) exists [18].*/ Suppes has shown how to express these axioms in terms of relations between probabilities [17]; the new axioms on probabilities (let us call them S) suffice to prove the existence of a utility function in the sense of Definition 1. The three conditions of Definition 2 are trivially equivalent to the continuity axioms of S. Finally, we have been able to prove that all the further axioms of S hold if the quadruple condition is satisfied (and provided of course the assumptions specified in I.2(a) and (b) hold). Hence we know that if the continuity and quadruple

---

*/ See also Franz Alt [1].

conditions of Definitions 2 and 3 hold, S holds, and there exists a utility function.

(d) A result similar to Theorem 4 was obtained by Debreu [4] under a different definition of stochastic continuity properties. Debreu has shown that there exists a utility function on A if the following conditions are satisfied:

(i) if a, b, c are in A and $P(b,a) \geq q \geq P(c,a)$ then there is a d in A such that $P(d,a) = q$;

(ii) the quadruple condition holds for A;

(iii) if x, y, z denote variable elements of A, then $P(x,z)$ depends continuously on $P(x,y)$ and $P(y,z)$.

We turn finally to the interesting case in which:

(e) the set A contains an unknown number (possibly finite) of alternatives. For this case no axiom system is known, and it has been conjectured by Scott and Suppes [16] that under certain natural restrictions on the form of axioms no axiomatization is possible.

It may be noted that in cases (b), (c), and (d) the axiom systems adequate to prove the existence of a utility function (in the sense of Definition 1) are adequate to prove also that any such function is unique up to a linear transformation (i.e., the existence of cardinal utilities).

We can submit to direct experimental test a set of the kind described in (a) containing a small, known, finite number of alternatives (let us call the set of alternatives under test T). If the quadruple condition is satisfied for every quadruple of alternatives in T, and T is a sample drawn from a larger set A, we may conclude--with a degree of confidence depending among other things on the size of the sample--that the quadruple condition holds for A. If our hypothesis is that a utility function on A exists then we need further information about A: for example we may know the (finite) number of its elements (case (a)), or we may hold that A is stochastically continuous (cases (c) and (d)).

In the experiment reported here, one hypothesis is that a utility function exists for the set consisting of all money wagers of a certain sort. If we can assume that A is stochastically continuous in the sense (c) or (d) (for example because the money amounts which enter the wagers are, approximately, continuous variables) and if, on the basis of our sample T, we have concluded that the quadruple condition holds for A, then we can conclude, by Theorem 4, that there exists a utility function on A.

Actually, we did not test for the quadruple condition on our sample T. Instead, we tested for certain implications of that condition: if T does not satisfy such an implication, we reject the hypothesis that T satisfies the quadruple condition. These implications involve triples (not quadruples) of alternatives and will be referred to as stochastic transitivity properties.

Even for relatively small finite sets of alternatives, the existence of a utility function in the sense of Definition 1 implies more than is implied by the quadruple condition alone. This fact suggests a view of the relation between experimental evidence and hypothesis which differs slightly from the one outlined in the preceding paragraphs. For each sample T drawn from A we may test all the conditions necessary and sufficient for the existence of a utility function on T (the general method is given above in the discussion of case (a)). We then consider confirmation of the existence of a utility function on T as inductive evidence for the existence of a utility function on A. As will be shown, the condition of strong stochastic transitivity about to be stated gives necessary and sufficient conditions for the existence of a utility function on a set consisting of three alternatives.

5. Definitions of stochastic transitivity. We say that weak stochastic transitivity holds in A if and only if, for all a, b and c in A,

5.1. if $P(a,b) \geq 1/2$ and $P(b,c) \geq 1/2$ then $P(a,c) \geq 1/2$. We say that strong stochastic transitivity holds in A if and only if, for all a, b and c in A,

5.2. if $P(a,b) \geq 1/2$ and $P(b,c) \geq 1/2$ then $P(a,c) \geq \max[P(a,b), P(b,c)]$.

These terms are due to S. Vail [19]. (We sometimes omit the word "stochastic.") Clearly 5.2 implies 5.1, but 5.1 does not imply 5.2; both are implied by the existence of a utility function and are therefore necessary conditions for the existence of such a function. 5.2 is equivalent to:

6. if $P(a,b) \geq 1/2$ then $P(a,c) \geq P(b,c)$.*/

Consider three fixed alternatives, a_1, a_2, a_3 and label the three relevant probabilities $P(a_1,a_2) = p_1$, $P(a_2,a_3) = p_2$, $P(a_3,a_1) = p_3$. The two kinds of transitivity condition applied to the set consisting of a_1, a_2, a_3 can then be expressed in the following symmetric form:

7.1. Weak transitivity: p_1, p_2, p_3 not all $> 1/2$ or $< 1/2$.

7.2. Strong transitivity:

$$p_1 \geq 1/2 \text{ if and only if } p_2 + p_3 \leq 1,$$

$$p_2 \geq 1/2 \text{ if and only if } p_3 + p_1 \leq 1,$$

$$p_3 \geq 1/2 \text{ if and only if } p_1 + p_2 \leq 1.$$

In the experiment reported in this paper, we are concerned with triples of alternatives. It is therefore interesting to note that if the set of alternatives consists of exactly three elements, a, b, c, then the condition of strong stochastic transitivity is not only necessary for the existence of a utility function (as mentioned at the end of the preceding sub-section (5)) but also sufficient. For, under strong transitivity, we may assume without loss of generality that $P(a,c) \geq P(a,b) \geq P(b,c) \geq 1/2$. The corresponding inequalities between utilities

*/ Proof: To show that 5.2 implies 6, assume $P(a,b) \geq 1/2$ and show that, by 5.2, $P(a,c) \geq P(b,c)$ for each of the three possible cases:
(1) $P(b,c) \geq 1/2$; then $P(a,c) \geq \max\,[P(a,b), P(b,c)] \geq P(b,c)$;
(2) $P(b,c) < 1/2 \leq P(a,c)$; then $P(a,c) \geq P(b,c)$;
(3) $P(b,c) < 1/2$, $P(a,c) < 1/2$; then $P(c,a) > 1/2$, hence $P(c,b) \geq \max\,[P(c,a), P(a,b)] \geq P(c,a)$, $P(a,c) > P(b,c)$. It may be left to the reader to prove the converse: that 6 implies 5.2.

(Definition 1) are: $u(a) - u(c) \geq u(a) - u(b) \geq u(b) - u(c) \geq 0$. These inequalities are satisfied, for example, by the following numbers: $u(a) = 1$, $u(c) = 0$, $u(b)$ = any number between, and including, 0 and 1.

In Table 1, the upper three cards show how we tested strong (and weak) transitivity experimentally. The subject made choices between the two columns on a card; the syllables on the left represent events determining the outcome of a wager. On the three cards there are altogether three alternatives (wagers) paired in each of the three possible ways. By testing whether condition 7.2 holds for a sample consisting of a number of such triples of alternatives, we obtain evidence for or against the hypothesis that a utility function exists on the set A from which the sample is drawn.

III. STOCHASTIC THEORY OF CHOICE BETWEEN SUBJECTIVELY EVEN-CHANCE WAGERS

In this section we deal with a special case of the stochastic theory of choice, exploiting some possible properties of choices between wagers of a special sort, namely those created by chance events with a "subjective probability of one half." The theoretical and experimental importance of the non-stochastic theory of choice for such wagers was first pointed out by Ramsey [14]; a formalization of the theory applied to finite sets, and reports of several experimental applications (including one with stochastic aspects) are given in Davidson, Suppes and Siegel [3].

We assumed in Section II that the set A of alternatives might contain wagers as well as sure outcomes; however, the formal developments made no use of this assumption.

Some additional primitive notions are needed.

1. A set X of <u>states of the world</u>. The subsets of X are called <u>events</u>, denoted by E, F, ... and forming a set $\mathcal{E}$.

2. If <u>a</u>, <u>b</u> are in A and E is in $\mathcal{E}$ then aEb is the wager which consists in getting <u>a</u> if E happens and getting <u>b</u> if E does not happen.

Definitions 3-6 of Section I are applicable to wagers; for example when P(aEb, cFd) = 1/2 we say that aEb and cFd are (stochastically) indifferent. We may presume that in certain cases absolute preference occurs. In particular, if $\underline{a}_1$, $\underline{a}_2$, $\underline{b}_1$, $\underline{b}_2$ are in A and $P(a_1,b_1) = 1 = P(a_2,b_2)$ then for any event E in $\mathcal{E}$, $P(a_1Ea_2, b_1Eb_2) = 1$.

<u>Definition</u> 3. An event E in $\mathcal{E}$ is an <u>even-chance event</u> if and only if, for every <u>a</u> and <u>b</u> in A,

$$P(aEb, bEa) = 1/2 \ .$$

If E is an even-chance event, we call aEb an <u>even-chance wager</u>. It is obvious that the notion of even-chance involved in this definition is <u>subjective</u>; it makes no appeal to the objective probability of E. The justification for our terminology is simple. Suppose a subject prefers <u>a</u> to <u>b</u>. If he thinks E is more likely to happen than not, he will choose aEb more often than bEa; if he thinks E less likely to happen than not, he will choose aEb less often than bEa. Hence he will choose aEb and bEa equally often if and only if he thinks E is as likely to happen as not, i.e., E has a "subjectively even chance."*/

*/ The next three pages (16-18) give, in the form of a footnote, an alternative reading of the original text (p. 19.1). Since Professor Marschak has not been able to review the contents of this footnote, he cannot be held responsible for it.

Definition 4. The subject is said to be unbiased if and only if, for any two even-chance events E and F in $\mathcal{E}$ and any a and b in A:

$$P(aEb, aFb) = 1/2 .$$

The chief concern of this section may be stated by giving a more restrictive version of Definition II.1:

Definition 5. A real valued function u is an even-chance wager utility function (or a utility function in the sense of Definition III.5) on A if and only if:

(a) u is a utility function on A in the sense of Definition II.1,

(b) for every a and b in A and E in $\mathcal{E}$, $u(aEb) = \frac{u(a)}{2} + \frac{u(b)}{2}$.

5(a) and (b) together express in stochastic form the usual hypothesis that a subject prefers the wager with the higher expected utility (applying this hypothesis to the case of even-chance wagers).

Now we wish to state conditions sufficient for the existence of an even-chance wager utility function. To this end we define the following condition:

Definition 6. The even-chance midpoint condition holds in A if and only if for every a and b in A and E in $\mathcal{E}$,

$$P(a, aEb) = P(aEb, b).$$

Definition 6 says that for any chance event E in $\mathcal{E}$, aEb is a stochastic midpoint between a and b (see Definition I.6). Next we show that if the quadruple condition (II.3) and the even-chance midpoint condition hold, then all events in $\mathcal{E}$ are even-chance events and the condition of unbiasedness (Definition 4) obtains. First we prove an elementary lemma which depends only on the quadruple condition.

Lemma 6.1. For all a, b, c and d in A, if $P(a,b) = P(b,d)$ and $P(a,c) = P(c,d)$ then $P(b,c) = 1/2$.

Proof: Suppose Lemma 6.1 were false, that is, its antecedent true and its consequent false. Then $P(b,c) \neq 1/2$, and hence, by the quadruple condition,

(1) $P(b,d) \neq P(c,d)$.

Then either $P(b,d) > P(c,d)$ or $P(c,d) > P(b,d)$. Assume, first, that $P(b,d) > P(c,d)$. Then, by the quadruple condition, we have

(2) $P(b,c) > 1/2$.

But from our assumption and the antecedent of the lemma, we have $P(a,b) > P(a,c)$ and hence, by the quadruple condition,

(3) $1/2 > P(b,c)$,

which contradicts (2). Assume, second, that $P(c,d) > P(b,d)$. Then, by the quadruple condition, we have,

(4) $P(c,b) > 1/2$.

But from our assumption and the antecedent of the lemma, we have $P(a,c) > P(a,b)$ and hence,

(5) $1/2 > P(c,b)$,

which contradicts (4). Therefore if the antecedent of the lemma is true, (1) is false, which proves the lemma.

Theorem 6.2. _If the quadruple and even-chance midpoint conditions hold in A, then every event E in $\mathcal{E}$ is an even-chance event._

Proof: For every E in $\mathcal{E}$, and _a_ and _b_ in A,

(1) P(b, bEa) = P(bEa, a) (even-chance midpoint condition)

(2) P(a, bEa) = P(bEa, b) ((1) and quadruple condition)

(3) P(a, aEb) = P(aEb, b) (even-chance midpoint condition)

(4) P(aEb, bEa) = 1/2 ((2), (3) and Lemma 6.1) Q.E.D.

Theorem 6.3. _If the quadruple and even-chance midpoint conditions hold in A, then the subject is unbiased._

Proof: By the even-chance midpoint condition we have, for any _a_ and _b_ in A and E and F in $\mathcal{E}$:

(1) P(a, aEb) = P(aEb, b) and

(2) P(a, aFb) = P(aFb, b).

Hence, using Lemma 6.1,

(3) P(aEb, aFb) = 1/2. Q.E.D.

Since the even-chance midpoint condition limits the wagers under consideration to even-chance wagers, we may, in what follows, simply write 'ab' for 'aEb.' (End of footnote)

Definition 4. The subject is said to be unbiased if and only if, for any two even-chance events E and F and any a and b in A: P(aEb, aFb) = 1/2. It is obvious that if this condition is satisfied and there exists a utility function u on A, then for any two even-chance events E and F and any a and b in A, u(aEb) = u(bEa) = u(aFb) = u(bFa). This justifies writing simply 'ab' for 'aEb' where E is any even-chance event; since we explicitly consider no other wagers, symbols for chance events need not enter our formalism.

The chief concern of this section may be stated by giving a more restrictive version of Definition II.1:

Definition 5. A real valued function u is an even-chance wager utility function (or a utility function in the sense of Definition III.5) on A if and only if:

(a) u is a utility function on A in the sense of Definition II.1,

(b) for every a and b in A and every even-chance event E,

$$u(aEb) = \frac{u(a)}{2} + \frac{u(b)}{2} .$$

3(a) and (b) together express in stochastic form the usual hypothesis that a subject prefers the wager with the higher expected utility (applying this hypothesis to the case of even-chance wagers). Clearly these conditions imply that u(aEb) has the same value for all even-chance events E in $\mathcal{E}$, and that the subject is unbiased.

Now we wish to state conditions sufficient for the existence of an even-chance wager utility function. To this end we define the following condition:

Definition 6. The even-chance midpoint condition holds in A if and only if the subject is unbiased and, for every a and b in A, P(a,ab) = P(ab,b). (Definition 6 says ab is a stochastic midpoint between a and b; see Def.I.6.)

We may now state a theorem analogous to II.4:

Theorem 7. If A is stochastically continuous then an even-chance wager utility function on A exists if and only if the quadruple condition (II.3) and the even-chance midpoint condition hold in A.

Proof: Suppose A is stochastically continuous. Then a function u on A such that $P(a,b) \geq P(c,d)$ if and only if $u(a) - u(b) \geq u(c) - u(d)$ exists if and only if the quadruple condition holds (Theorem II.4). Hence the quadruple condition is a necessary condition for the existence of an even-chance wager utility function. And if the quadruple condition is satisfied then a utility function in the sense of Definition II.1 exists; hence:

$P(a,ab) = P(ab,b)$ if and only if $u(a) - u(ab) = u(ab) - u(b)$,

that is,

$$u(ab) = \frac{u(a)}{2} + \frac{u(b)}{2} \quad .$$

Therefore the quadruple and even-chance midpoint conditions together provide necessary and sufficient conditions that a utility function in the sense of Definition III.3 exist, provided A is stochastically continuous.

An alternative statement of sufficient conditions may now be considered. We define:

Definition 8. The even-chance quadruple condition holds in A if and only if, for every a, b, c and d in A:

$$P(a,b) \geq P(c,d) \text{ if and only if } P(a,bc) \geq P(bc,d);$$

and assert:

Theorem 9. *If A is stochastically continuous then a utility function on A in the sense of Definition III.5 exists if and only if the even-chance quadruple condition holds in A.*[*]

We now establish an interesting consequence of the even-chance quadruple condition:

Theorem 10. *If the even-chance quadruple condition holds in A then for all a, b, c and d in A*

$$P(a,b) \geq P(c,d) \text{ if and only if } P(ad,bc) \geq 1/2.$$

We establish Theorem 10 by noting that if the even-chance quadruple condition holds, then $P(a,b) > P(c,d)$ is equivalent to $P(a,bc) \geq P(bc,d)$, which in turn is equivalent to $P(ad,bc) \geq P(bc,ad)$.

---

[*] Proof: It follows directly from Definition III.5 that if a utility function in the sense of that definition exists the even-chance quadruple condition holds. We prove the sufficiency of the even-chance quadruple condition by showing that it implies both the even-chance midpoint condition and the quadruple condition, and then applying Theorem 7. By the even-chance quadruple condition we have (replacing '*b*' by '*a*', and '*c*' and '*d*' by '*b*'):

$$(1) \quad P(a,a) = P(b,b) \text{ if and only if } P(a,ab) = P(ab,b).$$

The right side of (1) (i.e. the even-chance midpoint condition) is true since the left side is true by I.2(b). Using the even-chance quadruple condition again and assumption I.2(a) the following steps lead to the quadruple condition:

$$(2) \quad P(a,b) \geq P(c,d) \text{ if and only if } P(d,bc) \geq P(bc,a)$$

$$(3) \quad P(d,bc) \geq P(bc,a) \text{ if and only if } P(a,c) \geq P(b,d).$$

We have seen (in the proof of Theorem 9) that the even-chance quadruple condition implies the quadruple as well as the even-chance midpoint condition. On the other hand, these two conditions in conjunction do not imply the even-chance quadruple condition since they do not imply its consequence stated in the conclusion of Theorem 10.*/ Theorem 10 thus states a strong principle. It interlocks, in effect, the utility scales obtained by comparing differences in utility by two separate methods.

11. Definition of stochastic transitivity for utility intervals. We say that weak stochastic transitivity for utility intervals holds in A if and only if, for a, b, c, d, e and f in A,

11.1. if $P(bf,de) \geq 1/2$ and $P(ae,cf) \geq 1/2$ then $P(ab,cd) > 1/2$.

We say that strong stochastic transitivity for utility intervals holds in A if and only if, for all a, b, c, d, e and f in A,

11.2. $P(bf,de) \geq 1/2$ if and only if $P(ab,cd) \geq P(ae,cf)$.

---

*/ To show this, suffice it to consider the inequalities:

$$P(a,b) > P(c,d) > P(bc,ad) > 1/2 > P(ad,bc) > P(d,c) > P(b,a),$$

which contradict the conclusion of Theorem 10. Yet they are consistent with the conjunction of the quadruple and the even-chance midpoint condition, for the only relations to which those conditions in conjunction can apply in the present case are (apart from trivial repetitions): $P(a,b) > P(ad,bc)$ and $P(c,d) > P(bc,ad)$. The former relation yields $P(a,ad) > P(b,bc)$ and hence $P(ad,d) > P(bc,c)$, $P(ad,bc) > P(d,c)$, consistent with the assumed chain of inequalities; the latter relation yields, by similar steps, $P(bc,ad) > P(b,a)$, also consistent with the assumed inequalities.

The analogy between the transitivity conditions for alternatives (II.5) and transitivity conditions for intervals (III.11) may be brought out as follows. If a utility function in the sense of Definition III.5 exists then 11.1 is equivalent to the statement (holding identically for any six numbers):

$$\text{if } [u(b) + u(f)] - [u(d) + u(e)] \geq 0 \text{ and}$$
$$[u(a) + u(e)] - [u(c) + u(f)] \geq 0 \text{ then}$$
$$[u(a) + u(b)] - [u(c) + u(d)] \geq 0 ,$$

and hence to:

12.1. if $u(b) - u(d) \geq u(e) - u(f)$ and $u(e) - u(f) \geq u(c) - u(a)$ then $u(b) - u(d) \geq u(c) - u(a)$.

Similarly, 11.2 is equivalent to:

12.2. $u(b) - u(d) \geq u(e) - u(f)$ if and only if $[u(b) - u(d)] - [u(c) - u(a)] \geq [u(e) - u(f)] - [u(c) - u(a)]$.

Now let the length of the utility interval $u(b) - u(d) = I$, $u(e) - u(f) = J$ and $u(c) - u(a) = K$. Then 12.1 and 12.2 become similar in form to II.5.1 and II.6:

13.1. if $I \geq J$ and $J \geq K$ then $I \geq K$

13.2. $I \geq J$ if and only if $I-K \geq J-K$.

Thus 11.1 and 11.2 may be interpreted as stating conditions on utility intervals analogous to the conditions stated by II.5.1 and 5.2 for alternatives (whether or not these alternatives happen to be wagers).

However, it should be emphasized that in testing the transitivity of intervals we must make use of wagers; while we did use wagers in testing the transitivity of alternatives, this is not essential to the theory.

From 12.1 and 12.2 it is clear that the transitivity conditions for utility intervals are necessary for the existence of a utility function in the sense of Definition III.5.*/ To obtain evidence whether such a function exists for a limited set of outcomes consisting of winning and losing small amounts of money, we tested certain implications of 11.1 and 11.2 for sextuples of outcomes which may be regarded as samples from the total set of outcomes. Let us designate six specific money outcomes $a_1, a_2, a_3, a_4, a_5, a_6$ arranged in ascending order by monetary value. For reasons given in the next section we considered the following probabilities only:

14. $p_1 = P(a_1a_4, a_2a_3)$; $p_2 = P(a_6a_3, a_5a_4)$; $p_3 = P(a_5a_2, a_6a_1)$.

For these three probabilities the implications of 11.1 and 11.2 are just:

15.1. Weak transitivity of utility intervals:

$$p_1, p_2, p_3 \text{ not all } > 1/2 \text{ or } < 1/2.$$

*/ It was conjectured by the authors that if the set of alternatives is stochastically continuous then the conjunction of weak transitivity of alternatives and of weak transitivity of intervals is necessary and sufficient for the existence of a utility function in the sense of Definition III.5. While the manuscript was in preparation the conjecture was proved by G. Debreu using his definition of stochastic continuity - see Section II case (d) above.

15.2. Strong transitivity of utility intervals:

$$p_1 \geq 1/2 \text{ if and only if } p_2 + p_3 \leq 1,$$

$$p_2 \geq 1/2 \text{ if and only if } p_3 + p_1 \leq 1,$$

$$p_3 \geq 1/2 \text{ if and only if } p_1 + p_2 \leq 1.$$

It will be noted that conditions 15.1 and 15.2 are identical with conditions II.7.1 and II.7.2 where, of course, the three relevant probabilities are differently defined. If a utility function for the six outcomes $a_1, \ldots, a_6$ exists, 15.1 and 15.2 will be satisfied; but the converse is not in general true. The existence of a utility function in the sense of Definition III.5 implies, even for six outcomes, more than is implied by the transitivity of intervals condition (for example, III.10 is implied by the existence of a utility function but not by the transitivity of intervals condition); and the transitivity of intervals condition alone implies more, for six outcomes, than is tested by checking the relations given in 15.1 and 15.2 with p_1, p_2, p_3 as defined above.

The second line of specimen cards in Table 1 illustrates the method used in testing 11.1 and 11.2. Before the three pairs of wagers on these cards were offered the subject it was verified that the chance events underlying the designed money-wagers were even-chance events in the sense of Definition 3.*/ This justified assuming that all wagers on the cards were (for the given subject) even-chance wagers; and therefore it could be tentatively assumed that whenever the subject chose a wager (a column of a card) he could be interpreted as comparing two utility intervals,

---

*/ A more precise statement of the procedure used will be given in the next section.

represented by the rows of the card. On the three cards illustrated there are three pairs of identical rows; they correspond to the intervals I, J and K in 13.1 and 13.2. (From 12.1 and 12.2 it is clear that interchange of rows or of columns in a given card does not matter nor the interchange of alternatives in one column.)

IV. EXPERIMENTAL DESIGN

The experiment to be described was designed to test the plausibility of the hypothesis that (for given individuals) there exists a utility function in the sense of Definition III.5 (and therefore in the sense of Definition II.1), defined over a set of alternatives consisting of winning and losing small amounts of money and of even-chance wagers constructed from the basic alternatives. The individuals were 17 students from an elementary logic class at Stanford University. The general hypothesis was tested by testing certain of its consequences: stochastic transitivity (weak and strong) of alternatives as applied to triples of alternatives (interpreted here as wagers); and stochastic transitivity (weak and strong) of utility intervals as applied to sextuples of alternatives.

The obvious way of testing a stochastic theory of choice is to estimate probabilities of choice from frequencies of choice observed when the subject is repeatedly offered the same alternatives. This method, common from psychophysical experiments, has been used with apparent success by a number of workers in decision theory. These workers (who include Mosteller and Nogee, Ward Edwards and Papandreou) were, of course, aware of the memory effect, and used various techniques in the attempt to cope with it. In a pilot study for the present experiment we found that with wagers of the sort we wished to use the subject almost always makes the same choice when offered the same pair of alternatives; thus we would be forced to estimate almost every probability as 0 or 1. The wagers between which the subjects had to choose had the same actuarial value; the wagers could therefore be assumed to be close in subjective utility. Remarks by the subjects led to the suspicion that the cause of the unforeseen consistency was the subject's ability to remember his previous choices (although various masking procedures were attempted such as reversing the order in which the wagers in a pair were offered, and inserting other offers between repetitions of the identical pair of wagers). In psychophysical experiments memory cannot have this effect since the subject is given no way of identifying the repetition of a stimulus.

Therefore in order to avoid the effect of memory, the same pair of wagers was never offered twice to a subject. The method used for testing our hypothesis under this restriction is explained in the next section.

Each subject was asked to make 319 choices; a choice consisted in a verbal response ('A' or 'B') to a stimulus-card of the kind illustrated in Table 1. In 107 cases selected (with certain limitations to be mentioned later) at random, and unknown in advance to the subject, the response of the subject was followed by playing off the wager selected, and the subject lost or won the appropriate amount of money.

The 319 stimulus-cards were designed as follows. Every card displayed four figures (positive or negative) representing a possible outcome consisting of losing or winning the amount of money shown. On the left were two nonsense syllables (WUH and XEQ; ZOJ and ZEJ; QUJ and QUG) which stood for chance events. The events were created by the subject tossing a die with one nonsense syllable on three faces and another nonsense syllable on the other three faces. In an effort to offset recency and memory effects three different dice were used. The two right hand columns, marked 'A' and 'B' represent the wagers between which the subject was to choose.

For testing hypotheses concerning the existence of an even-chance wager utility function it was necessary to ascertain whether the events created by the three dice just described were even-chance events. In practice this was tested indirectly by assuming, for any money amounts m and n (in cents):

$$P(mEn,\ nEm) = 1/2 \text{ if and only if } P(mEn,\ n\text{-}1\ E\ m) > 1/2 \text{ and } P(mEn,\ n\text{+}1\ E\ m) < 1/2.$$

Previous experiments had shown that, given this modified interpretation, subjects generally accept the nonsense syllable dice as generating even-chance events; therefore we tested each die only a few times with each subject (see Davidson, Suppes and Siegel [3], p. 56 and Table 1, p. 57). In all, 12 stimulus-cards were used to test the dice; three additional cards were added to this group to familiarize the subject with other sorts of choice.

The remaining cards were intended to test the transitivity of alternatives and of intervals (II.5 and III.11). Thirty-eight sequences of seven money amounts were chosen such that the money amounts, in the light of previous experiments (Mosteller and Nogee; Davidson, Suppes and Siegel), would be approximately evenly spaced in utility for most subjects. Table 2 gives the first 19 sequences; the other 19 sequences were produced from the first by reversing the signs (thus wins become losses and vice versa). This symmetry provides a simple guarantee that the actuarial value of the complete set of wagers is zero; why this is desirable will be explained below. Eight cards were made for each 7-tuple of money amounts, yielding 8x38 = 304 cards in all. Using the letters at the top of Table 2, the eight cards showed the following patterns:

1	2	3	4	5	6	7	8
a b	c d	a b	b c	a c	d e	b c	b d
d c	f e	f e	e d	f d	g b	g f	g e

It will be observed that the triads 3, 4, 5 and 2, 7, 8 each contain just three alternatives (wagers, represented as columns) and thus may be used to test the transitivity of alternatives. Triads 1, 2, 3 and 4, 6, 7 each compare, in effect, three intervals (represented by rows) and thus may be used to test the transitivity of intervals. Because of this overlap between triples we have achieved some economy in the number of observations: the total of 304 cards yields 76 triples designed to test the transitivity of alternatives and 76 triples designed to test the transitivity of intervals.

The assignment of one of the three dice to a specific card was random. Because certain wagers (not cards) were repeated once, the column (A or B) on which a wager appeared was randomized; the row (top or bottom) assigned to an outcome in a wager was also randomized. Finally, the order in which the cards appeared was randomized, except that the fifteen cards used for learning and to test that the dice created even-chance events preceded all others, and all three cards from a given triple appeared during the same session.

We may now make explicit two rules employed in limiting the offers appearing on the stimulus cards. No sure thing alternatives were allowed on the ground that these might distort the results should there exist a specific utility or disutility of gambling. The second rule is intended to eliminate cases of absolute preference. In any given triple of cards there are six distinct outcomes. Let us assign the numbers 1, 2, ..., 6 to the six outcomes in order of monetary value; the number assigned an outcome denotes its rank. The rule is this: on any given card, the sum of the ranks of the two outcomes in one wager must be equal to the sum of the ranks of the two outcomes in the other wager (see III.14). Since the outcomes are chosen to be approximately evenly spaced in utility, the rule is designed to insure that two wagers which are compared shall not differ too strongly in expected utility. In application, no two wagers on one card differed by more than $4\frac{1}{2}$ cents in actuarial value. When the transitivity of intervals is tested for six outcomes the two rules just mentioned limit the pairs of wagers to be compared to exactly three.

Of the 17 subjects 6 were women and 11 were men. Subjects were tested individually. Each subject came to three sessions, spaced a few days apart; two sessions were never on the same day for a given subject, nor more than five days apart. A session lasted between 35 and 55 minutes. Subjects were asked not to discuss the experiment during the testing; none of them had any detailed knowledge of game theory or decision theory.

At the beginning of the first session a subject was shown the three nonsense syllable dice and the game he was to play was explained. The subject was given $2.00 credit (in chips) and told that this was his stake for the three sessions. At the end of the three sessions, his chips would be redeemed in cash; if he had won, he would receive $2.00 plus his winnings; losses would come out of the $2.00; greater losses would have to be paid out of his own pocket.

The first 15 stimulus cards have been described; of these, 12 tested whether the subject accepted the dice as creating even-chance events, and three were for learning purposes. In effect, every subject did accept the dice as 'fair.' All of the first 15 cards were played off: after the subject gave his response by choosing wager A or B, he put the indicated die in a leather cup, shook, and rolled. Depending on the outcome, the experimenter then collected from or paid out to the subject the appropriate number of chips. The rest of the first session consisted in responses to 88 more stimulus cards testing the two sorts of transitivity. Of these, 25 choices were played off; the subject did not, of course, know whether a card would be played until after he had made his choice. The cumulative expected win for a subject who always chose the wager with the higher actuarial value was +44¢ for the 25 cards which were played off. Subjects were urged to take as long as they wished to make a decision.

During the second session the subject was asked to make 112 decisions; of these, 36 were played off. The cumulative expected win for the actuarial chooser was +39¢.

The last session called for 104 decisions of which 31 with an expected win of +55¢ were played off. During this last session the experimenter could play off additional wagers to increase the winnings of an unlucky subject.

As mentioned above, the actuarial value of the total of all wagers offered was zero. Since the wagers between which a subject was to choose seldom had exactly the same actuarial value, a consistent "actuarial chooser" could have expected to win if every choice had been played off. The cards chosen for playing had a small positive actuarial value for the "random chooser" and a higher actuarial value for the actuarial chooser. The hope was that the average subject with average luck would slowly increase the sum at his disposal; its size would not vary enough to influence choices substantially. It may be doubted whether this hope was entirely realized. In any case for many subjects, the sum at their disposal changed fairly radically during the play, and verbal comments by subjects suggested that this influenced choices. The highest total win (for all three sessions) was $4.87 (including the original $2.00 stake); the least fortunate subject received a few cents less than $2.00. However several subjects had their winnings 'artificially' increased during the last session by the experimenter naming for playoff certain cards on which both wagers had high positive actuarial value; unknown to the subject, it had been decided in advance that no subject would average less than $1.00 an hour for his time.

V. STATISTICAL DECISION RULES

Transitivity regions. Consider the three related probabilities p_1, p_2, p_3 as defined for the statement of conditions II.7.1 and II.7.2, or as defined for the statement of conditions III.15.1 and III.15.2. Let us denote by $p^i = < p_1^i, p_2^i, p_3^i >$ the i^{th} ordered triple of probabilities so defined. p^i is a point in the unit cube U, since each component of p is between 0 and 1.

We now define two sub-regions of U:

Region W: p^i obeys condition II.7.1 (or III.15.1) (region of weak transitivity).

Region S: p^i obeys condition II.7.2 (or III.15.2) (region of strong transitivity).

Obviously region S is included in region W. The hypothesis pairs to be tested may be stated:

Hypothesis H_w: For all i, p^i is in W;

Hypothesis H_w^o: There exists an i such that p^i is in U-W;

Hypothesis H_s: For all i, p^i is in S;

Hypothesis H_s^o: There exists an i such that p^i is in U-S.

Note that each of the hypotheses has two empirical interpretations; one concerns stochastic transitivity of alternatives, the other stochastic transitivity of utility intervals. We need not distinguish between the two interpretations in discussing the method of statistical testing.

Since a given choice was presented to a subject only once, it was impossible to estimate the probability triples p^i from observed frequencies. Corresponding to a given p^i we made one observation consisting of the three responses of a subject to a triple of related stimulus cards. Suppose, for the sake of simplicity of exposition, that the pairs of wagers on a related triple of stimulus cards are arranged in the order suggested by the definitions of p_1, p_2, p_3 in Sections II and III (this has been done for the triples of cards shown in Table 1). Then if the subject chooses column A on the first card, there is greater likelihood that $p_1 > 1/2$ than that $p_1 < 1/2$; if he chooses column A on the third card, there is greater likelihood that $p_3 > 1/2$ than that $p_3 < 1/2$. An <u>observation</u> is an ordered triple of responses; there are just eight possible observations:

$$O_1 = \langle A,A,A \rangle \qquad O_5 = \langle B,A,A \rangle$$

$$O_2 = \langle A,A,B \rangle \qquad O_6 = \langle B,A,B \rangle$$

$$O_3 = \langle A,B,A \rangle \qquad O_7 = \langle B,B,A \rangle$$

$$O_4 = \langle A,B,B \rangle \qquad O_8 = \langle B,B,B \rangle$$

In a non-stochastic theory observations O_1 and O_8 would be cases of intransitivity. In a stochastic theory they merely strengthen the evidence in favor of (stochastic) intransitivity. To avoid confusion we call such observations <u>cyclical</u> because, e.g., $\langle A,A,A \rangle$ means that a certain wager <u>a</u> was chosen in preference to <u>b</u>; <u>b</u> to <u>c</u>; and <u>c</u> in preference to <u>a</u>, thus forming a cycle.

In its strict formulation, our problem is analogous to the following simpler (one- instead of three-dimensional) problem: "Test the hypothesis that each coin made by a certain coin-making machine has a bias, not necessarily an equally strong one for all coins, in favor of falling heads. You are permitted to take a finite number of coins and to toss each coin just once." Each coin of this example corresponds to a triple of choices from three pairs of our wagers. The parameter-space is, respectively, the unit-interval (0,1) or the unit-cube U. The interval (1/2, 1) which contains the probability of a biased coin falling heads corresponds to our transitivity region W (or S) which contains all probability-triples if the subject satisfies the transitivity condition. Should this formulation be accepted, then, out of the infinite set of potential observations (coins, triples of choices) it would suffice for a single one to be outside of a specified region (the bias-interval for coins, the transitivity region for response-triples), in order to rule out the hypothesis in question. But such a fact cannot be ascertained empirically, from a finite number of observations. The problem becomes accessible to empirical test if it is reformulated as follows: "A coin-making machine is characterized by an unknown probability distribution of the chance variable p (probability of a coin falling heads); one is permitted to toss coins, each only once, in order to get evidence about the distribution of p." The chance variable p corresponds, in the theory of stochastic choice, to the triple: $< p_1, p_2, p_3 >$ defined above.

For example, one might test the following hypothesis about the distribution of p: the proportion of coins (or of triples of wagers) whose p falls into a specified region is at least 95%. This approach has been used in a later study, by H. D. Block and J. Marschak [2]; with regions like S, this statistical problem is rather complicated. In the present study, we chose a simpler though more arbitrary approach by adding the following assumption: p is uniformly distributed about an unknown region which is either the whole space of possible p's (the unit cube, in our case) or a specified region (such as W or S). We have thus two pairs of alternative hypotheses:

$$\begin{cases} H_W: & p^i \text{ is distributed uniformly over } W, \text{ and } \operatorname{Prob}(p^i \in W) = 1 \\ H_o: & p^i \text{ is distributed uniformly over } U, \text{ and } \operatorname{Prob}(p^i \in U) = 1 \end{cases}$$

$$\begin{cases} H_s: & p^i \text{ is distributed uniformly over } S, \text{ and } \operatorname{Prob}(p^i \in S) = 1 \\ H_o: & p^i \text{ is distributed uniformly over } U, \text{ and } \operatorname{Prob}(p^i \in U) = 1. \end{cases}$$

It turns out that for testing the statistically reformulated hypotheses, all that matters (the "sufficient statistic") is the number of cyclical observations. Computations yield the following probabilities of a cyclical observation:

	Probability of a Cyclical Observation
If H_o is true	$\frac{20}{80} = 25.00$ %
If H_W is true	$\frac{15}{80} = 18.75$ %
If H_s is true	$\frac{11}{80} = 13.75$ %

The reasoning leading to the first figure (25%) is obvious: if p^i is distributed uniformly over the unit cube, each of the eight possible observations $O_1, \ldots, O_8$ is equi-probable. Since two of these are cyclical, the probability that a given observation is cyclical if H_o is true is 1/4. The other two figures (15/80 and 11/80) were obtained*/ by integrating over the specified region (W or S, respectively) the expression

$$p_1 p_2 p_3 + (1 - p_1)(1 - p_2)(1 - p_3),$$

(i.e., the probability that O_1 or O_8 will occur), and dividing by the volume of that region.

The <u>decision rule</u> used was this (we state it for H_w; that for H_s is analogous): Accept H_w if the number r of cyclical observations is less than c, where c (a number obtainable from tables of binomial distribution) is such that Prob ($r < c$, when H_o is true) = Prob ($r \geq c$, when H_w is true). The two probabilities just written are, respectively, that of committing the error of accepting H_w when H_o is true, and that of committing the error of accepting H_o when H_w is true. By making them equal, we make the larger of the two error probabilities a minimum. The significance level of test is equal to both of them. If instead of this "minimax" principle that takes into account both kinds of error we had followed the custom of merely "testing the null hypothesis," the decision could have been made, given our number of observations, with a much lower

---

*/ With the help of Karol Valpreda Walsh and Robert C. Mercer, which we gratefully acknowledge.

probability of errors of Type I (called "significance level"). The latter would then have been defined as the probability of rejecting H_o when it is true. However, the nature of the problem forces us to treat the null hypothesis and its alternative symmetrically.

In the present experiment we made on each subject a total of 76 observations for each hypothesis; 26 observations for each hypothesis were made during the last session. Applying our computations and decision rule to these figures we obtain:

	Number of Observations	
	n = 76 (all sessions)	n = 26 (Session III)
Decision rules		
Accept H_w if r is less than:	17	6
Significance level:	24%	35%
Accept H_s if r is less than:	15	5
Significance level:	10%	23%

One obtains, of course, much lower significance levels if one is permitted to regard the responses of all subjects as belonging to the same statistical population, thus increasing the sample size by the factor of 17(the number of subjects). Using the normal distribution formula, one can compute the decision rules at which the probabilities of errors of both kinds are equated; these rules are, for samples of this magnitude: reject H_w if and only if the proportion of cyclical observations is larger than 21.7 percent; reject H_s if and only if this proportion is larger than 18.7 percent (compare these figures with the last line of Table 3). The significance levels are as follows: for Session III, .055 for H_W and .001 for H_S; for the total of all sessions, .003 for H_W and negligible for H_S.

Finally, we observe that if the proportion of cyclical observations falls very much below the probability indicated above for weak or strong transitivity, we shall presume that the assumption of uniform distribution over the (weak or strong) transitivity region is to be corrected: we shall have to assign lower weights to those points of the region that lie near its boundaries other than the facets of the unit cube.

As another and possibly preferable way of exploiting the information more fully, Herman Chernoff and Roy Radner*/ suggested computing the likelihood ratio for each pair of hypotheses and each possible set of observations. Let $P_O(r,n)$ be the probability that, out of n responses, exactly r are cyclical if H_O is true. With a corresponding notation for the cases when H_W or H_S are true, we can use the ratios

$$L_W(r,n) = P_W(r,n)/P_O(r,n) \quad \text{and} \quad L_S(r,n) = P_S(r,n)/P_O(r,n)$$

to convey the confidence that one may attach, on the basis of observations, to the hypothesis H_W or H_S, each against the hypothesis H_O. Using the binomial distribution formulas, one derives from the probabilities of cyclical observations: (20/80 for H_O, 15/80 for H_W and 11/80 for H_S), the two likelihood ratios

$$L_W(r,n) = (9/13)^r \cdot (13/12)^n \quad \text{and} \quad L_S(r,n) = (11/23)^r \cdot (23/20)^n .$$

---

*/ In an oral communication

These numbers are tabulated below, for n = 76 and n = 26 (with r ranging from 0 to, respectively, 17 and 7), and can be used to interpret, subject by subject, the results shown on Table 3 for the total of observations and for Session III.*/

Likelihood ratios (approximate)

No. of cyclical responses: r =	H_w against H_o n = 76	H_w against H_o n = 26	H_s against H_o n = 76	H_s against H_o n = 26
0	430	8.0	40 000	38
1	300	5.5	20 000	18
2	210	3.8	9 300	8.7
3	140	2.7	4 500	4.1
4	100	1.8	2 100	2.0
5	69	1.3	1 000	0.9
6	49	0.9	490	0.5
7	33	0.6	240	0.2
8	23	0.4	110	0.1
9	16		54	
10	11		26	
11	7.6		12	
12	5.3		5.9	
13	3.6		2.8	
14	2.5		1.3	
15	1.7		0.6	
16	1.2		0.3	
17	0.8		0.1	

*/ Again, if we regard all subjects as belonging to the same statistical population, the likelihood ratios are naturally much higher: they are of the order of many thousands, for both H_w and H_s, even if the proportion of cyclical observations is as high as 15 percent.

VI. RESULTS OF EXPERIMENT

Table 3 summarizes the main experimental findings. For a large majority of subjects, the number of cyclical responses falls far below the expected frequency under strong transitivity (both of alternatives and of intervals). Under the null hypothesis, one out of four triples could, on the average, be expected to be cyclical; this would result in 38 cyclical responses for the total of 152 triples, or 19 for the 76 triples which tested each of the two varieties of transitivity. The highest total number of cyclical responses for any subject was 28 (Subject J), while Subject F had the highest number of cyclical responses (17) for a set of 76 triples. The expected numbers of cyclical responses in 76 triples for H_w and H_s are 14.25 and 10.45 respectively. Two subjects (J and N) in the case of the transitivity of alternatives, and two subjects (C and F) in the case of the transitivity of intervals, exceeded the number of cyclical responses expected on the assumption that H_w is true. The average number of cyclical responses for all subjects on the 76 triples testing transitivity of alternatives is 10.4 while the average on the triples testing transitivity of intervals is 10.7. These figures are close to the prediction of 10.45 cyclical responses if strong transitivity holds.

The last line of Table 3 should be related to what was said in Section 5 about using the responses of all subjects as a sample from the same population. On this basis (and under the assumptions of Section 5 regarding a priori distribution) weak as well as strong transitivity, for both alternatives and intervals, should be accepted at a quite low (i.e., strict) significance level).*/

Table 4 applies the Decision Rules of Section V to each subject (using the results from all sessions). For all subjects, Hypothesis H_w of weak transitivity had to be accepted with respect both to alternatives and to intervals (though at a very modest, i.e., high, significance level: 24%). For all but two subjects, strong transitivity had to be accepted (at a significance level of 10%), but the two subjects were not the same for alternatives and for intervals. The correlation of .49 between behavior with respect to alternatives and behavior with respect to intervals is just significant at the 5% level for 17 observation-pairs; however, it should not be judged significant at that level if one considers the fact that some of the data were used in two ways as described in Section IV.

An interesting feature of the data displayed in Table 3 is the change in the proportion of cyclical observations from session to session. While not always evident for individual subjects, when the results for all subjects are averaged there is a systematic decrease in the percentage of cyclical observations from session to session both for alternatives and for intervals. During the first session the overall proportion of cycles is 13.4%; during the second session 10.5%; during the third 7.4%.

---

*/ If weak transitivity is accepted for both alternatives and intervals and if the set of even-chance money-wagers can be regarded as stochastically continuous then--as remarked in a footnote in Sec. III--a utility function in the sense of Definition III.5 exists. The size of the sample used for the joint test of transitivity of alternatives and of intervals is somewhat reduced by the overlapping between the sets of cards used for these two tests separately (see Sec. IV); this raises the significance level somewhat.

For the third session, the Decision Rule indicates that the null hypothesis must be accepted for one subject (N) for transitivity of alternatives, and for one subject (J) for transitivity of intervals. There is also one subject in each category (Subjects J and C respectively) for whom weak but not strong transitivity must be accepted. In every other case the hypothesis of strong transitivity may be accepted.

During the three sessions there was an increase in the correlation between performance with respect to alternatives and performance with respect to intervals, as follows:

Session	I	II	III	Total
Correlation	.00	.52	.58	.49

The change from I to II is significant at the 5% level.

VII. DISCUSSION

Assuming a utility function of money unique up to a linear transformation exists for a subject, it is possible to make some rough inferences from the data concerning the shape of the utility curve (a curve which plots utility against the money amount of the basic alternatives). Table 5 classifies the responses of each subject to the 304 cards used to test transitivity. On 249 cards the two wagers had different actuarial values; the second and fifth columns in Table 5 show in how many cases the subject chose the wager with the higher actuarial value ('actuarial responses') and in how many cases the subject chose the wager with the lower actuarial value ('counter-actuarial responses'). On every card, one wager involved both the highest and the lowest money amount, while the other wager involved outcomes of intermediate money value; see III.14. Responses may therefore be classified according as the subjects chose the wager with the greater or the lesser dispersion (the dispersion of an even-chance wager is the difference in money value between the two outcomes). The second and third columns of figures in Table 5 show how often the counter-actuarial choices favored high or low dispersion. 55 cards showed wagers of equal actuarial value; these responses are classified to indicate whether the high or the low dispersion wager was chosen. In evaluating the figures in Table 5 it is necessary to know that of the 249 cards with wagers of unequal actuarial value, 124 paired the higher actuarial value with the higher dispersion and 125 paired the higher actuarial value with the lower dispersion.

A subject for whom utility was linear in money and with absolute preferences between all pairs of wagers offered would always choose the wager with the higher actuarial value. There is no subject who did this. Subjects A and O come closest with 22 and 33 counter-actuarial answers. Both these subjects show some preference for low dispersion when they depart from the actuarial choice.

We may call a subject conservative who, when departing from the actuarial answer, more often than not chose the wager with the lower dispersion, for such a subject would, from an insurance point of view, be paying for the privilege of taking the smaller risk. In the same way we may call a subject venturesome who, when departing from the actuarial answer, more often than not chose the wager with the higher dispersion. Using this criterion, five subjects were conservative and twelve were venturesome. In the light of previous experimental results with college students (vid. Mosteller and Nogee [11] and Davidson, Suppes and Siegel [3]), the percentage of venturesome subjects is perhaps surprising. Part of the explanation may lie in the fact that the subjects were volunteers and knew, before they volunteered, that the experiment involved some financial risk.

A subject who invariably chose the higher or the lower dispersion would yield no cyclical observations (no pair of wagers showed the same dispersion as this would lead to absolute preference), while the consistent actuarial chooser would be certain to show no more than six cyclical triads of responses (there were six triples of cards where each pair of wagers had the same actuarial value). It is therefore an interesting question to what extent the results obtained were due to actuarial, conservative or venturesome strategies (conscious or otherwise) on the part of the subjects. Table 5 makes it obvious that no subject consistently followed any of the three policies. Comparison of Tables 3 and 5 brings out the fact that the four subjects with the largest number of cyclical observations (Subjects C, D, F and J) include none of the five subjects with the largest number of counter-actuarial responses. On the other hand, Subject L, with the lowest number (4) of cyclical triples, seems to owe this score in part to a conservative taste for low dispersion wagers, while Subject A, with only five cyclical triples, has the fewest counter-actuarial responses.

We may also ask whether the frequency of cyclical observations depends on the size of the differences in the actuarial values of each pair of wagers. One would expect cyclical observations to occur relatively more often in cases where these differences are small and thus provide no guidance, or only a weak one, for the choice between two wagers and the ranking of three wagers. In Table 6, four groups--A, B, C, D--of triples are defined, the number of observations (i.e., the number of triples times 17, the number of subjects) is entered for each group, and the frequency of cyclical vs. non-cyclical observations given. A finer grouping of triples was precluded by sample size limitations. In fact, the distinction between Groups A and B proved to be statistically insignificant (and hence the unexpectedly lower percentage of cyclical observations in Group A compared with Group B is not statistically significant). Significant and interesting results are obtained by comparing Group (A,B) (which is the union of A and B and thus consists of all triples with actuarial differences not exceeding 1/2 cent on any card) either with Group C (actuarial differences not smaller than 1 cent on any card) or with the composite group (C,D)(actuarial difference exceeds 1/2 cent on at least one card). These comparisons tend to confirm the hypothesis that small actuarial differences favor the occurrence of cyclical choices.

It may be asked whether the decrease in the proportion of cyclical observations from session to session was accompanied by an increase in the proportion of actuarial responses. The following figures show that it was.

	Sessions		
	I	II	III
Number of cards with wagers of different actuarial value	1241	1598	1394
Number of actuarial choices	810	1106	979
(Percentage)	(65.3 %)	(69.2 %)	(70.2 %)
Number of triples	748	952	884
Number of non-cyclical observations	648	845	819
(Percentage)	(86.6 %)	(89.5 %)	(92.6 %)

The increase in the proportion of actuarial choices was statistically significant ($P < .05$) from Session I to Session II and definitely not significant ($P > .50$) from Session II to Session III. On the other hand, the proportion of non-cyclical observations increased highly significantly ($P < .01$) from Session II to Session III, although it did not increase significantly ($P > .2$) from Session I to Session II. This suggests that, to the extent to which there was an increase in non-cyclical observations it should not be explained by the increase in actuarial choices. However, a more detailed analysis would be necessary to clarify this point.

In the table just given, the proportion of non-cyclical observations to the total number of triples of cards is, in each of the three sessions, consistently higher than the proportion of actuarial choices to the total number of cards. A formal test shows the difference between the proportions to be statistically highly significant. Thus, considering all subjects to belong to the same population, the probability that a subject's response to a triad of cards will be non-cyclical is higher than the probability that his response to a single card will be actuarial.

Broadly speaking the present experiment shows that in its context decisions are better explained by certain implications of a stochastic decision theory than by the assumption that choices are made at random. The interest of this result would be impugned if the same data could be as well or better explained by alternative hypotheses. For this reason we have been considering the plausibility of the claim that, to the extent that subjects avoided cyclical triads of responses (and hence tended to verify the hypotheses under test), this was due to the more or less consistent employment of actuarial, venturesome or conservative policies by the subjects. In Table 7 we attempt a direct comparison of three alternative theories as predictors of certain observed choices.

Assume that a related triple of cards was always arranged in the "normal form" shown in Table 1 and that the three cards were presented to the subject in order from left to right. If a subject chose one left and one right column for the first two cards, the hypotheses H_W (or H_S) as well as H_O would predict equal probabilities of choice for each wager on the third card. The interesting cases arise when the left column or the right column is chosen on both the first two cards; then the hypotheses H_W and H_S will give the higher probability of choice to the opposite column on the third card (for example, if column A is chosen on each of the first two cards, column B is more apt to be chosen on the third card if the hypotheses are true).

The first two columns of figures in Table 7 show how often the subject chose the more probable wager on the third card for those cases where both of the first two choices were wager A or wager B. The third and fourth columns show how often, for the same cards for which results are given in the first two columns, the subject chose the wager with the higher actuarial value. The total number of predictions is lower because some cards had two wagers with equal actuarial values. The last two columns show, for the same cards again, how often the subject chose in accord with his general tendency to favor low or high dispersion wagers. This tendency was determined from the figures in Table 5 showing how often the counter-actuarial responses of the subject favored low or high dispersion.

At the bottom of Table 7 the three theories are compared with respect to the percentage of correct predictions. The expected utility theory is superior to the other two with 81.6% correct predictions; the actuarial theory is slightly better than the dispersion theory with 72.2% correct predictions as compared to 69.0% correct predictions. Each of the differences between these proportions (taken pairwise) is statistically significant ($P < .01$). It is also worth noting that for every individual subject the expected utility theory predicted better than the dispersion theory; however the actuarial theory was very slightly superior to the expected utility theory for two subjects (Subjects C and J).

A remark is called for concerning the marked decrease in the proportion of cyclical observations from session to session. It seems attractive to call this a learning phenomenon. However, what was learned was certainly not connected with specific cards, wagers, money amounts, or triples of cards, for none of these was repeated in two sessions (specific cards and triples were not repeated even in the same session). The evidence indicates that the subjective probability of the chance events was firmly established from the start. We may say that the subject learned transitive behavior or that he learned to maximize expected utility (at least in the sense of our stochastic definitions of these terms). There is also evidence that some subjects learned to make actuarial choices. The matter seems worthy of fuller study, both theoretical and experimental.

VIII. SUMMARY

The experiment was designed to test whether certain conditions hold which are necessary for the existence of a utility function over the set of money wagers as well as over the set of money amounts. Seventeen subjects were tested individually in three sessions each. Every choice involved a risk since subjects did not know in advance which choices would be played off, that is, would result in actually paying (or receiving) money. The main results may be summarized as follows:

1. For all subjects, the number of intransitive triads of responses (called 'cyclical observations') was, as required by the hypotheses, less than the number expected by chance.

2. A statistical test, based on responses to non-repeated choice situations, indicates the acceptance, for all subjects, of the hypotheses of weak stochastic transitivity of alternatives and weak stochastic transitivity of utility intervals.

3. For fifteen subjects in each case, the acceptance of the hypotheses of strong stochastic transitivity of alternatives and of intervals is indicated.

4. Both in testing the transitivity of alternatives and the transitivity of intervals, there was a systematic decrease in the number of cyclical responses from session to session.

5. A comparison shows the superior accuracy of a stochastic theory of decision in predicting certain choices as compared to two alternative theories. On the whole the evidence does not appear to support the claim that the low number of cyclical observations can be wholly explained by simple policies based on the actuarial values of wagers, or on the degree of risk (dispersion).

TABLE 1

Specimen Stimulus-Cards

For testing transitivity of alternatives:

	A	B
ZOJ	-5¢	+36¢
ZEJ	-21¢	-38¢

	A	B
QUG	+36¢	-54¢
QUJ	-38¢	+22¢

	A	B
WUH	-54¢	- 5¢
XEQ	+22¢	-21¢

For testing transitivity of utility intervals:

	A	B
ZOJ	- 6¢	+ 5¢
ZEJ	+24¢	+13¢

	A	B
QUG	+38¢	+31¢
QUJ	+13¢	+24¢

	A	B
WUH	+31¢	+38¢
XEQ	+ 5¢	- 6¢

TABLE 2

Money Amounts (Wins and Losses) Used in

Constructing Stimulus-Cards

Sequence	a	b	c	d	e	f	g
1	-17	-12	- 5	+ 2	+ 8	+17	+21
2	-20	-16	-11	- 4	+ 3	+10	+15
3	-12	- 5	+ 2	+ 9	+14	+19	+23
4	-22	-18	-14	- 9	- 2	+ 5	+12
5	-14	- 8	- 4	+ 5	+13	+20	+26
6	-36	-17	- 8	+ 2	+10	+21	+27
7	-13	- 8	- 5	- 1	+ 4	+ 7	+12
8	-35	-28	-22	-15	- 8	- 2	+ 6
9	-27	-16	- 4	+ 7	+21	+34	+47
10	- 6	+ 1	+ 8	+15	+22	+30	+34
11	-39	-25	- 9	+ 6	+23	+40	+56
12	-37	-22	- 6	+ 5	+21	+38	+54
13	- 7	- 4	- 1	+ 3	+ 6	+ 8	+11
14	-14	- 9	- 6	- 2	+ 3	+ 6	+11
15	- 6	+ 2	+12	+21	+29	+36	+42
16	-24	-20	-14	- 7	+ 1	+ 8	+13
17	-31	-17	- 5	+ 8	+21	+33	+46
18	-11	- 3	+ 4	+12	+19	+24	+28
19	-17	- 6	+ 5	+13	+24	+31	+38

TABLE 3

Number of Cyclical Observations for each Subject

	Testing Transitivity of Alternatives				Testing Transitivity of Intervals			
Sessions:	I	II	III	Total	I	II	III	Total
Number of triples offered	22	28	26	76	22	28	26	76
Expected number under uniform distribution over the region of:	Number of Cyclical Observations							
unit cube	5.50	7.00	6.50	19.00	5.50	7.00	6.50	19.00
weak transitivity	4.13	5.25	4.88	14.25	4.13	5.25	4.88	14.25
strong transitivity	3.03	3.85	3.58	10.45	3.03	3.85	3.58	10.45
Subject								
A	4	0	0	4	0	0	1	1
B	3	5	2	10	1	6	3	10
C	5	3	3	11	6	5	5	16
D	4	7	0	11	4	5	2	11
E	1	0	0	1	0	3	3	6
F	3	6	0	9	5	10	2	17
G	2	1	2	5	3	0	0	3
H	2	1	1	4	4	2	1	7
I	1	2	1	4	3	3	0	6
J	2	9	5	16	3	3	6	12
K	4	2	2	8	3	5	3	11
L	1	1	0	2	1	1	0	2
M	2	2	1	5	7	5	1	13
N	6	3	7	16	2	0	3	5
O	4	2	1	7	3	2	0	5
P	1	3	3	7	2	1	3	6
Q	7	5	2	14	1	4	2	7
Average no. of cyclical observations:	3.06	3.06	1.76	7.88	2.82	3.24	2.06	8.12
Percentage of cyclical observations:	13.9	10.9	6.8	10.4	12.8	11.6	7.9	10.7

TABLE 4

All Sessions (76 Triads); Distribution of Subjects by the Number of Cyclical Observations. (Correlation Coefficient = .49)

No. of cyclical observations for options (rows); No. of cyclical observations for intervals (columns); No. of subjects (last column)

No. of cyclical observations for options:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	No. of subjects:
1						E												1
2		L																1
3																		0
4	A					I	H											3
5			G										M					2
6																		0
7					O	P												2
8											K							1
9																	F	1
10										B								1
11											D					C		2
12																		0
13																		0
14							Q											1
15																		0
16					N							J						2
17																		0
No. of subjects:	1	1	1	0	2	3	2	0	0	1	2	1	1	0	0	1	1	

Rows 1–14: Accept Strong and Weak Transitivity; rows 15–17: Accept Weak Transitivity Only.

Columns 1–13: Accept Strong and Weak Transitivity; columns 14–17: Accept Weak Transitivity Only.

TABLE 5

Responses for each Subject According to Actuarial Value and Dispersion

(Total number of wagers = 304)

Subject	Pairs of wagers of unequal actuarial value				Pairs of wagers of equal actuarial value		All pairs of wagers	
	Actuarial responses	Counter-actuarial responses						
		Dispersion chosen		Total	Dispersion chosen		Dispersion chosen	
		High	Low		High	Low	High	Low
A	227	10	12	22	11	44	134	170
B	149	75	25	100	41	14	213	91
C	162	54	33	87	37	18	184	120
D	165	75	9	84	43	12	234	70
E	174	45	30	75	29	26	168	135
F	161	73	15	88	41	14	223	81
G	149	79	21	100	43	12	225	79
H	175	65	9	74	46	9	225	79
I	200	36	13	49	34	21	181	123
J	162	50	37	87	27	28	164	140
K	167	54	28	82	34	21	184	120
L	182	15	52	67	18	37	103	201
M	136	49	64	113	28	27	137	167
N	160	53	36	89	37	18	177	127
O	216	12	21	33	12	43	127	177
P	165	23	61	84	18	37	104	200
Q	145	58	46	104	34	21	169	135

TABLE 6

Actuarial Characteristics of Triples Offered and the Frequency of Cyclical Observations

Character of triples offered (d = difference between the actuarial values of the two wagers on a card.)	Number of observations: Total	Cyclical	Non-cyclical	Percentage of cyclical in total
A. $d = 0$ on all three cards	102	17	85	16.7
B. $0 < d \leq 1/2¢$ on all three cards	272	59	213	21.7
(A,B). $0 \leq d \leq 1/2¢$ on all three cards	(374)	(76)	(298)	(20.3)
C. $d \geq 1¢$ on all three cards	442	49	393	11.1
D. All other triples	1768	147	1621	8.3
(C,D). $d > 1/2¢$ on at least one of the three cards	(2210)	(196)	(2014)	(9.3)
TOTAL	2584	272	2312	10.5

Variation of frequency between groups	χ^2	P	Significant?
A against B	1.3	> .25	No
C against D	3.2	> .05	Hardly
(A,B) against C	13.7	< .005	Yes
(A,B) against (C,D)	43	< .005	Yes

TABLE 7

Comparison of Expected-Utility, Actuarial, and

Dispersion Theories as Predictors

Subject	Expected Utility Theory		Actuarial Theory		Dispersion Theory	
	Correct predictions	Incorrect	Correct	Incorrect	Correct	Incorrect
A	84	5	74	5	56	33
B	63	20	39	29	60	23
C	46	27	42	20	44	29
D	86	22	69	24	84	24
E	76	7	56	19	50	33
F	80	26	68	28	74	32
G	86	8	47	33	80	14
H	100	11	74	22	92	19
I	71	10	66	8	50	31
J	49	28	43	25	42	35
K	52	20	43	17	48	25
L	83	4	60	17	64	23
M	57	18	43	25	50	25
N	63	21	53	22	59	25
O	76	12	69	8	57	31
P	70	13	52	24	68	15
Q	70	21	41	35	48	43
Totals	1213	273	939	361	1026	460
Percentage Correct	81.6		72.2		69.0	

References

[1] Alt, Franz, "Über die Messbarkeit des Nutzens," Zeitschrift für Nationaloekonomie, 1936, Vol. 7, No. 2, pp. 161-169.

[2] Block, H. D. and Jacob Marschak, "Random Orderings," Cowles Foundation Discussion Paper No. 42, 1957 (mimeographed).

[3] Davidson, Donald, Patrick Suppes and Sidney Siegel, Decision Making, An Experimental Approach, Stanford, 1957.

[4] Debreu, Gerard, "Stochastic Choice and Cardinal Utility," Econometrica, 1958 (forthcoming).

[5] Edwards, Ward, Articles on Probability Preferences, American Journal of Psychology, 1953, 66, pp. 345-364; 1954, 67, pp. 56-67; 1954, 67, pp. 68-95.

[6] Fechner, Gustav Theodor, Elemente der Psychophysik, 1859.

[7] Guilford, J. P., Psychometric Methods, New York, 1954.

[8] Luce, Duncan, "A Probabilistic Theory of Utility," Technical Report No. 14 (mimeographed), 1957.

[9] Luce, Duncan and Howard Raiffa, Games and Decisions, New York, 1957 (Appendix I, "A Probabilistic Theory of Utility," pp. 371-384).

[10] Marschak, Jacob, "Norms and Habits of Decision Making Under Uncertainty," Mathematical Models of Human Behavior - Proceedings of a Symposium, Stamford, Connecticut, 1955.

[11] Mosteller, F. and P. Nogee, "An Experimental Measurement of Utility," Journal of Political Economy, LIX (1951), pp. 371-404.

[12] von Neumann, J. and O. Morgenstern, Theory of Games and Economic Behavior, 2nd. ed. Princeton, New Jersey, 1947.

[13] Papandreou, Andreas G., with the collaboration of O. H. Sauerlander, O. H. Brownlee, L. Hurwicz, and W. Franklyn, "A Test of a Stochastic Theory of Choice," University of California Publications in Economics, Vol. 16, No. 1, pp. 1-18.

[14] Ramsey, Frank P., The Foundations of Mathematics and Other Logical Essays, London, 1931.

[15] Savage, L. J., Foundations of Statistics, New York, 1951.

[16] Scott, Dana and P. Suppes, "Foundational Aspects of Theories of Measurement," Journal of Symbolic Logic, Vol. 23, 1958.

[17] Suppes, P., "A Set of Axioms for Paired Comparisons" (dittoed) Center for Behavioral Sciences, 1956.

[18] Suppes, P., and M. Winet, "An Axiomatization of Utility Based on the Notion of Utility Differences," Management Science, I (1955) pp. 259-270.

[19] Vail, Stephan, "A Stochastic Model for Utilities," Seminar on the Application of Mathematics to the Social Sciences, University of Michigan, 1953 (dittoed).

STANFORD UNIVERSITY

Technical Reports Distribution List

Contract Nonr 225(17)

(NR 171-034)

Armed Services Technical Information Agency Arlington Hall Station Arlington 12, Virginia	5
Commanding Officer Office of Naval Research Branch Office Navy No. 100, Fleet Post Office New York, New York	35
Director, Naval Research Laboratory Attn: Technical Information Officer Washington 25, D. C.	6
Office of Naval Research Group Psychology Branch Code 452 Department of the Navy Washington 25, D. C.	5
Office of Naval Research Branch Office 346 Broadway New York 13, New York	1
Office of Naval Research Branch Office 1000 Geary Street San Francisco 9, Calif.	1
Office of Naval Research Branch Office 1030 Green Street Pasadena 1, California	1
Office of Naval Research Branch Office Tenth Floor The John Crerar Library Building 86 East Randolph Street Chicago 1, Illinois	1
Office of Naval Research Logistics Branch, Code 436 Department of the Navy Washington 25, D. C.	1
Office of Naval Research Mathematics Division, Code 430 Department of the Navy Washington 25, D. C.	1
Operations Research Office 6935 Arlington Road Bethesda 14, Maryland Attn: The Library	1
Office of Technical Services Department of Commerce Washington 25, D. C.	1
The Logistics Research Project The George Washington University 707 - 22nd Street, N.W. Washington 7, D. C.	1
The RAND Corporation 1700 Main Street Santa Monica, Calif. Attn: Dr. John Kennedy	1
Library Cowles Foundation for Research in Economics Box 2125 Yale Station New Haven, Connecticut	1
Center for Philosophy of Science University of Minnesota Minneapolis 14, Minnesota	1

Institut für Math. Logik
Universitat
Schlossplatz 2
Münster in Westfalen
Germany 1

Professor Ernest Adams
Department of Philosophy
University of California
Berkeley 4, California 1

Professor Maurice Allais
15 Rue des Gates-Ceps
Saint-Cloud, (S.-O.)
France 1

Professor Norman H. Anderson
Department of Psychology
Yale University
333 Cedar Street
New Haven, Connecticut 1

Professor T. W. Anderson
Center for Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, California 1

Professor K. J. Arrow
Department of Economics
Stanford University
Stanford, California 1

Professor Richard C. Atkinson
Department of Psychology
University of California
Los Angeles 24, California 1

Dr. R. F. Bales
Department of Social Relations
Harvard University
Cambridge, Massachusetts 1

Professor Alex Bavelas
Department of Psychology
Stanford University 1

Professor Gustav Bergman
Department of Philosophy
State University of Iowa
Iowa City, Iowa 1

Professor E. W. Beth
Bern, Zweerskade 23, I
Amsterdam, Z.,
The Netherlands 1

Professor Max Black
Department of Philosophy
Cornell University
Ithaca, New York 1

Professor David Blackwell
Department of Statistics
University of California
Berkeley 4, California 1

Mr. Gordon Bower
Department of Psychology
Yale University
New Haven, Connecticut 1

Professor R. B. Braithwaite
King's College
Cambridge, England 1

Professor C. J. Burke
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor R. R. Bush
The New York School of Social Work
Columbia University
2 East Ninety-first Street
New York 28, New York 1

Dr. Donald Campbell
Department of Psychology
Northwestern University
Evanston, Illinois 1

Professor Rudolf Carnap
Department of Philosophy
U.C.L.A.
Los Angeles 24, California 1

Professor C. West Churchman
School of Business Administration
University of California
Berkeley 4, California 1

Dr. Clyde H. Coombs
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. Gerard Debreu
Cowles Commission
Box 2125
Yale Station
New Haven, Conn. 1

Dr. Mort Deutsch
Bell Telephone Laboratories
Murray Hill, New Jersey 1

Professor Robert Dorfman
Department of Economics
Harvard University
Cambridge 38, Massachusetts 1

Dr. Ward Edwards
Lackland Field Unit No. 1
Operator Laboratory
Air Force Personnel and Training
Research Center
San Antonio, Texas 1

Dr. Jean Engler
Institute of Statistics
University of North Carolina
Chapel Hill, North Carolina 1

Professor W. K. Estes
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor Robert Fagot
Department of Psychology
University of Oregon
Eugene, Oregon 1

Dr. Leon Festinger
Department of Psychology
Stanford University 1

Professor M. Flood
Willow Run Laboratories
Ypsilanti, Michigan 1

Professor Maurice Fréchet
Institut H. Poincaré
11 Rue P. Curie
Paris 5, France 1

Dr. Milton Friedman
Center for Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, Calif. 1

Dr. Eugene Galanter
Department of Psychology
University of Pennsylvania
Philadelphia 4, Pa. 1

Dr. Murray Gerstenhaber
University of Pennsylvania
Philadelphia, Pennsylvania 1

Dr. I. J. Good
25 Scott House
Cheltenham, England 1

Dr. Leo A. Goodman
Statistical Research Center
University of Chicago
Chicago 37, Illinois 1

Professor Nelson Goodman
Department of Philosophy
University of Pennsylvania
Philadelphia, Pa. 1

Professor Harold Gulliksen
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 1

Professor Louis Guttman
Israel Institute of Applied
Social Research
David Hamlech No. 1
Jerusalem, Israel 1

Dr. T. T. ten Have
Social - Paed. Instituut
Singel 453
Amsterdam, Netherlands 1

Professor Carl G. Hempel
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Dr. Ian P. Howard
Department of Psychology
University of Durham
7, Kepier Terrace
Gilesgate
Durham, England 1

Professor Leonid Hurwicz
School of Business
University of Minnesota
Minneapolis 14, Minn. 1

Professor Lyle V. Jones
Department of Psychology
University of North Carolina
Chapel Hill, North Carolina 1

Professor Donald Kalish
Department of Philosophy
University of California
Los Angeles 24, California 1

Dr. Leo Katz
Department of Mathematics
Michigan State College
East Lansing, Michigan 1

Professor John G. Kemeny
Department of Mathematics
Dartmouth College
Hanover, New Hampshire 1

Professor T. C. Koopmans
Cowles Foundation for Research
in Economics
Box 2125, Yale Station
New Haven, Connecticut 1

Professor W. Kruskal
Department of Statistics
Eckart Hall 127
University of Chicago
Chicago 37, Illinois 1

Dr. David La Berge
Department of Psychology
University of Indiana
Bloomington, Indiana 1

Professor Douglas Lawrence
Department of Psychology
Stanford University 1

Dr. Duncan Luce
Department of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Dr. W. G. Madow
Engineering Research Division
Stanford Research Institute
Menlo Park, California 1

Professor Jacob Marschak
Box 2125 Yale Station
New Haven, Connecticut 1

Dr. Samuel Messick
Educational Testing Service
Princeton University
Princeton, New Jersey 1

Professor G. A. Miller
Department of Psychology
Harvard University
Cambridge 38, Massachusetts 1

Dr. O. K. Moore
Department of Sociology
Box 1965
Yale Station
New Haven, Connecticut 1

Professor Sidney Morgenbesser
Department of Philosophy
Columbia University
New York 27, New York 1

Professor Oskar Morgenstern
Department of Economics and
Social Institutions
Princeton University
Princeton, New Jersey 1

Professor Frederick Mosteller
Department of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Professor Ernest Nagel
Department of Philosophy
Columbia University
New York 27, New York 1

Dr. Theodore M. Newcomb
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Professor A. G. Papandreou
Department of Economics
University of California
Berkeley 4, California 1

Dr. Hilary Putnam
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Professor Willard V. Quine
Department of Philosophy
Emerson Hall
Harvard University
Cambridge 38, Massachusetts 1

Professor Roy Radner
Department of Economics
University of California
Berkeley 4, California 1

Professor Howard Raiffa
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor Nicholas Rashevsky
University of Chicago
Chicago 37, Illinois 1

Dr. Frank Restle
Department of Psychology
Michigan State University
East Lansing, Michigan 1

Professor David Rosenblatt
American University
Washington 6, D. C. 1

Professor Alan J. Rowe
Management Sciences Research
Project
University of California
Los Angeles 24, California 1

Professor Herman Rubin
Department of Mathematics
University of Oregon
Eugene, Oregon 1

Dr. I. Richard Savage
School of Business
University of Minnesota
Minneapolis, Minn. 1

Professor L. J. Savage
Committee on Statistics
University of Chicago
Chicago, Illinois 1

Mr. Dana Scott
Department of Mathematics
Princeton University
Princeton, New Jersey 1

Dr. C. P. Seitz
Special Devices Center
Office of Naval Research
Sands Point
Port Washington
Long Island, New York 1

Dr. Marvin Shaw
School of Industrial Management
Massachusetts Institute of
Technology
50 Memorial Drive
Cambridge 39, Massachusetts 1

Dr. Sidney Siegel
Center for Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, California 1

Professor Herbert Simon
Carnegie Institute of Technology
Schenley Park
Pittsburgh, Pennsylvania 1

Dr. Herbert Solomon
Teachers College
Columbia University
New York, New York 1

Professor K. W. Spence
Psychology Department
State University of Iowa
Iowa City, Iowa 1

Dr. F. F. Stephan
Box 337
Princeton University
Princeton, New Jersey 1

Mr. Saul Sternberg
Department of Social Relations
Emerson Hall
Harvard University
Cambridge 38, Massachusetts 1

Professor S. Smith Stevens
Memorial Hall
Harvard University
Cambridge 38, Massachusetts 1

Dr. Donald W. Stilson
Department of Psychology
University of Colorado
Boulder, Colorado 1

Dr. Dewey B. Stuit
108 Schadffer Hall
State University of Iowa
Iowa City, Iowa 1

Professor Alfred Tarski
Department of Mathematics
University of California
Berkeley 4, California 1

Professor G. L. Thompson
Department of Mathematics
Dartmouth College
Hanover, New Hampshire 1

Dr. Robert L. Thorndike
Teachers College
Columbia University
New York, New York 1

Professor R. M. Thrall
University of Michigan
Engineering Research Institute
Ann Arbor, Michigan 1

Dr. Masanao Toda
Department of Experimental Psychology
Faculty of Letters
Hokkaido University
Sapporo, Hokkaido, Japan 1

Dr. E. Paul Torrance
Bureau of Educational Research
University of Minnesota
Minneapolis, Minn. 1

Professor A. W. Tucker
Department of Mathematics
Princeton University, Fine Hall
Princeton, New Jersey 1

Dr. Ledyard R. Tucker
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 1

Professor Edward L. Walker
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Professor Morton White
Department of Philosophy
Harvard University
Cambridge 38, Massachusetts 1

Dr. John T. Wilson
National Science Foundation
1520 H Street, N. W.
Washington 25, D. C. 1

Professor Kellog Wilson
Department of Psychology
Duke University
Durham, North Carolina 1

Professor J. Wolfowitz
Department of Mathematics
Cornell University
Ithaca, New York 1

Professor O.L. Zangwill
Psychology Laboratory
Downing Place
Cambridge, England 1

Professor Alan Ross Anderson
Department of Philosophy
Yale University
New Haven, Connecticut 1

Dr. Juliette Popper
Department of Psychology
Indiana University
Bloomington, Indiana 1

Mr. R.L. Shuey
General Electric Research Lab
Schenectady, N. Y. 1

Professor Robert McGinnis
Dept. of Sociology
University of Wisconsin
Madison 6, Wisc. 1

Professor Lyle E. Bourne, Jr.
Dept. of Psychology
University of Utah
Salt Lake City, Utah 1

Additional copies for project
leader and assistants and
reserve for future requirements 25