JOURNAL OF MATHEMATICAL PSYCHOLOGY 13, 323–337 (1976)

# Tests of the Betweenness Property of Expected Utility*

CLYDE H. COOMBS AND LILY C. HUANG

*Department of Psychology, University of Michigan, Ann Arbor, Michigan 48104*

In order that maximizing some kind of an expectation be acceptable as a descriptive theory of risky decision making it is necessary that a gamble made up of a probability mixture of two others lie between them in the perference order. It should not be the most preferred nor the least preferred of the three. The two experiments reported here test that condition and find it is significantly violated. According to Portfolio theory and expected risk theory the mixture may be most preferred but not least preferred. This condition is found to obtain.

## 1. INTRODUCTION

The literature in decision theory is widely varied in source, intent, and content. It comes from economics, sociology, psychology, statistics, and mathematics (see for example, Luce and Raiffa, (1957), and Lee, (1971)), and from engineering and management sciences (see for example, Starr, 1969). The intent of some is normative, some descriptive, and some is action oriented (see for example, Dunn, (1972)). Throughout this literature a dominant role is played by the concept of maximizing expected utility[1] so it never ceases to be probed both experimentally and theoretically and from both a normative and descriptive point of view. Because of the predominance of expected utility theory, the introduction of a notion of risk and risk-preference has not been as intensively pursued as it might have. Hence serious consideration of any alternative theory of decision making is best motivated by doubt thrown on maximizing expected utility.

* This research was supported by National Science Foundation Grant GB-15653 to the University of Michigan. We wish to thank Teri Berger and Robert Judson for assistance in the experiments and David H. Krantz, Steven Shevell, Paul Slovic, and J. E. Keith Smith for useful comments on the analysis and compositions, and Amos Tversky for substantial help and stimulating discussions.

[1] If $X$ is an option with mutually exclusive and exhaustive outcomes $(x_1, ..., x_i, ..., x_n)$ with associated probabilities $(p_1, ..., p_i, ..., p_n)$ and $u_b$ are real valued functions defined on the outcomes, then $EU(X) = \sum_{i=1}^{n} p_i u_i(x_b)$. If the outcomes are numbers, like points or money, then $X$ is referred to as a gamble and only one utility function is involved. If the $p_b$ are not known in the sense of objective probabilities but instead correspond to the subjective likelihood of the occurrence of events, $E_i$, then $S(E_i)$ replaces $p_b$ and the theory is called maximizing subjective expected utility (SEU Theory).

Here we examine a betweenness property of probability mixtures which is a necessary condition for the maximization of expected utility. We show that another theory of risky decision making, called Portfolio Theory, does not require this condition. Then we report two experiments which reveal significant violation of this property and hence the maximization of expected utility but not of Portfolio Theory.

We will present the betweenness property for the case of two-outcome gambles as its generalization to $n$-outcome gambles is direct and uncomplicated.

Let $A$ be a gamble with outcomes $a_1$ and $a_2$ with probabilities $p$ and $1 - p$ respectively and $C$ be a gamble with outcomes $c_1$ and $c_2$ with probabilities $q$ and $1 - q$ respectively. The expected utility of such a gamble as $A$ is $EU(A) = pU(a_1) + (1 - p)\, U(a_2)$. Maximizing expected utility requires that

$$A \succcurlyeq C \qquad \text{iff} \quad EU(A) \geqslant EU(C)$$

where $\succcurlyeq$ is a binary preference relation and $\geqslant$ is the natural order on the real numbers.

A probability mixture of $A$ and $C$, designated $(A, t, C)$ means that $A$ is played with probability $t$ otherwise $C$. So $(A, t, C)$ is a gamble with outcomes $a_1$ with probability $pt$, $a_2$ with probability $(1 - p)t$, $c_1$ with probability $q(1 - t)$, and $c_2$ with probability $(1 - q)(1 - t)$.

The expected utility of $(A, t, C)$ is

$$EU(A, t, C) = pt\, U(a_1) + (1 - p)t\, U(a_2) + q(1 - t)\, U(c_1) + (1 - q)(1 - t)\, U(c_2),$$

which reduces to

$$EU(A, t, C) = t\, EU(A) + (1 - t)\, EU(C).$$

It follows readily that

$$EU(A) \geqslant EU(C) \qquad \text{iff} \quad EU(A) \geqslant EU(A, t, C) \geqslant EU(C)$$

and hence we have the following betweenness property introduced as an axiom by von Neumann and Morganstern (1953, p. 26, 3 : B : a and 3 : B : b):

$$A \succcurlyeq C \qquad \text{iff} \quad A \succcurlyeq (A, t, C) \succcurlyeq C.$$

Under expected utility theory, therefore, a probability mixture of two gambles lies between them in preference. That is, the mixture of two gambles cannot be preferred to both of them nor can both gambles be preferred to their mixture. The betweenness property is also a requirement of SEU Theory because subjective probabilities of complementary events must add to one and lie in the unit interval, so the tests and results reported here are equally relevant to both theories.

If we let $A$ and $C$ designate two gambles with the same expected value and let $B$ designate a probability mixture of them then only the preference orderings $ABC$ and

*CBA* are admissible under expected utility theory. These two orderings will be called monotone orderings. The other four possible orderings violate EU Theory. These four are divided into two classes with two orderings in each, one of these classes is compatible with Portfolio Theory and the other is treated as reflecting an error rate in a manner to be described.

Portfolio Theory (Coombs, 1969; Coombs & Huang, 1970), is a risk-preference theory but leaves the notion of risk undefined. The notion of risk is itself an object of theoretical and experimental research; see, for example, Huang (1971a,b), Pollatsek and Tversky (1970), and Coombs and Huang (1970). The details of Portfolio Theory are not important here except for the assumption it makes that if several gambles have the same expected value but differ in risk, then an individual has a single-peaked preference function[2] over them. We add to this Huang's result (Huang, 1971a,b; see also Coombs & Bowen, 1971) that risk satisfies the betweenness property for gambles which have the same expected value, i.e., the probability mixture of two gambles which differ in risk but have the same expected value will have an intermediate level of risk. From these two assumptions, that for two gambles with the same expected value a probability mixture of them has an intermediate level of risk and that an individual has a single-peaked preference function over risk, then it follows immediately that for some $t \in [0, 1]$ he will prefer the probability mixture $B = (A, t, C)$ over both component gambles $A$ and $C$ if one of them is more risky than his optimal level of risk and the other is less risky. So the two preference orderings *BAC* and *BCA*, which are violations of EU Theory, are compatible with Portfolio Theory and are called *strictly folded* orderings. The two monotone orderings, *ABC* and *CBA*, are compatible with Portfolio Theory also, as well as EU Theory.

The remaining two logically possible orderings, *ACB* and *CAB*, are ones in which the gamble $B$, a probability mixture of $A$ and $C$ is the least preferred. These orderings could be described as resulting from a single-peaked *avoidance* function, an inverted preference function, so we will refer to them as *inverted* orderings. These inverted orderings violate EU Theory, Portfolio Theory, and, indeed, are inadmssible under any theory seriously proposed as a general theory of risky decision making. The frequency of occurrence of inverted orderings is used as an estimate of random error.

The two experiments are direct tests of the betweenness property in that a pair of gambles with the same expected value but presumably differing in risk are used to construct a triple of gambles in which the third is a probability mixture of the other two. For example, if $A$ were the gamble to toss for one dollar and $C$ the gamble to toss for five dollars, we could construct a third gamble, $B$, by formin a 50/50 mixture of $A$ and $C$ which would have outcomes (—\$5, —\$1, \$1, \$5) each with probability one-fourth. According to Huang's theory of expected risk, $B$ would have a level of

[2] A single-peaked function over $x$ may be defined as follows: if $a \succcurlyeq b \succcurlyeq c$ with respect to an attribute $x$ such as risk, then a single-peaked preference function over $x$ allows any permutation in which $b \succcurlyeq a$ or $b \succcurlyeq c$ where $\succcurlyeq$ is a binary preference relation.

risk intermediate between $A$ and $C$ and according to Portfolio Theory could be preferred to both of the others, a strictly folded preference ordering. Only the monotone preference orderings $ABC$ and $CBA$ are admissible under EU Theory. The inverted preference orderings, $ACB$ and $CAB$ will be used to evaluate the significance of the number of strictly folded orderings violating EU Theory. Strictly folded orderings are dual to inverted orderings and should therefore have the same rate of occurrence due to random error.

The two experiments differ in the way in which preference orders on triples were obtained and in the way in which the gambles in a triple differed from each other. In the first experiment the subject rank orders each triple directly and the gambles differed in outcomes. In the second experiment the subject replicated pair comparison preferences and the gambles in a triple all had the same outcomes and differed only in their respective probabilities.

## 2. Experiment I

### *Subjects*

The subjects were 26 student volunteers paid for participation in the experiment.

### *Stimuli*

The gambles used in the first session were of the form win $x$ with probability $p$ otherwise $y$, $(x, p, y)$, and are presented in Table 1. There are three sets of nine gambles each, the sets differing in expectation. Set I has gambles with $E = -20¢$; Set II has gambles with $E = 0$; and Set III has gambles with $E = 20¢$. The gambles in Sets I and III were constructed from those in Set II by subtracting or adding 20¢, respectively, to all outcomes. One gamble not used in the analysis was added to each set of games (set in brackets in Table 1) to make a total of ten gambles so a balanced incomplete block design could be constructed for each set which provided 15 blocks of six gambles each which contain five replications on every pair. Over all three sets, then, there were $15 \times 3 = 45$ blocks of six gambles each.

All gambles were represented by a picture of a spinner board without the spinner. The probabilities were represented by appropriate sections of the spinner board with winning and losing amounts indicated numerically. Two real spinner boards were constructed for practice purpose and real playing.

### *Procedure*

Each subject participated in two sessions, one week apart.

Session 1: In the first session the subjects were first familiarized with the gambles by playing some for practice. They were told that they would be allowed to play one game

TABLE 1

Gambles Used as Stimuli[a] in Experiment I

| | | | | | | |
|---|---|---|---|---|---|---|
| Set I $E = -20¢$ | $8.80 | 0.1 | −$1.20 | $0.80 | 0.9 | −$9.20 |
| | $2.13 | 0.3 | −$1.20 | $0.80 | 0.7 | −$2.53 |
| | $0.80 | 0.5 | −$1.20 | [$1.20 | 0.5 | −$0.80] |
| | $0.23 | 0.7 | −$1.20 | $0.80 | 0.3 | −$0.63 |
| | −$0.09 | 0.9 | −$1.20 | $0.80 | 0.1 | −$0.31 |
| Set II $E = 0$ | $9.00 | 0.1 | −$1.00 | $1.00 | 0.9 | −$9.00 |
| | $2.33 | 0.3 | −$1.00 | $1.00 | 0.7 | −$2.33 |
| | $1.00 | 0.5 | −$1.00 | [$0.80 | 0.5 | −$1.20] |
| | $0.43 | 0.7 | −$1.00 | $1.00 | 0.3 | −$0.43 |
| | $0.11 | 0.9 | −$1.00 | $1.00 | 0.1 | −$0.11 |
| Set III $E = +20¢$ | $9.20 | 0.1 | −$0.80 | $1.20 | 0.9 | −$8.80 |
| | $2.53 | 0.3 | −$0.80 | $1.20 | 0.7 | −$2.13 |
| | $1.20 | 0.5 | −$0.80 | [$1.00 | 0.5 | −$1.00] |
| | $0.63 | 0.7 | −$0.80 | $1.20 | 0.3 | −$0.23 |
| | $0.31 | 0.9 | −$0.80 | $1.20 | 0.1 | $0.09 |

[a] The gambles in brackets were included to augment each set so $n = 10$ and provide a convenient bibd. They were not used in the analysis.

for real at the end of the experiment and that the game would be selected by the experimenter on the basis of choices made by the subject during the experiment.

The 45 blocks of stimuli were arranged in random order. A subject was given one block of stimuli at a time. He was asked to indicate his preference order on the six games as if he would have to play one of them once. There was a 15 minute break after the subject had ordered 15 blocks.

Only Set III with $E = 20¢$ was used for the construction of the triples to be presented in the second sessions. The other two sets had been included in the first session for another purpose, to make some comparisons relevant to Portfolio Theory, and no further use of that data is reported here.

Preparation for Session 2: Between Sessions 1 and 2 the data from the first session on Set III were analyzed for each individual and the triples constructed which were to to be used in the second session. Each individual's data were used to construct a rank order of his preferences over the nine stimuli in Set III. This was done by counting the number of times each gamble was chosen over the others in the set and this total vote count yielded the rank order, the higher the total vote count for a gamble, the higher it was in the rank order.

The rank ordered I scale obtained in this manner was then divided in half, alternately providing two sections, one consisting of the five odd numbered ones in the preference order and the other section consisting of the four even numbered games in the preference order. A 50/50 probability mixture was then constructed for pairs of games in which one was from one section and one from the other.[3] The design may be represented by a matrix of five rows and four columns with each of the 20 cells containing a 50/50 probability mixture of the corresponding row and column—the row and column gambles being two-outcome games and the mixture being a three or four-outcome game. So a triple of gambles consisted of a row gamble, a column gamble, and the corresponding cell gamble constructed from the mixture. The set of 20 triples was prepared for each subject separately because it depended on the rank order of the nine games in his I scale. An example of a triple for one subject is the following:

$$A = (\$2.53, 0.3, -\$0.80)$$
$$C = (\$1.20, 0.1, \$0.09)$$
$$(A, 1/2, C) = (\$2.53, 0.15; \$1.20, 0.05; \$0.09, 0.45; -\$0.80, 0.35).$$

In this instance the gambles $A$ and $C$ differed in both amount to win and amount to lose so their mixture is a four-outcome gamble.

Session 2: In the second session, each subject indicated his preference order for the three games in each triple under the instructions to assume he was to play one of the games once; only transitive response patterns were possible in this session.

## Results

### *Measures of Consistency and Intransitivity*

Every pair of stimuli in each set was replicated five times, so a subject's choices on each pair were split either 5 — 0, 4 — 1, or 3 — 2. There are 36 pairs in each set so each subject's average split on all three sets is reported in the first column of Table 2. The subjects are numbered and ordered in table from most to least consistent on the basis of this overall average. The subjects are also divided in two halves in the table, the more and the less consistent.

If an individual were responding by chance to each pair with a choice probability of one-half, the expected value and standard deviation for the mean split on 108 pairs would be 3.44 and 0.059, respectively, so a mean split of 3.57 or more would indicate a significant degree (0.01 level) of non-random choice. The lowest overall level of consistency evinced by any subject exceeds this level substantially.

Transitivity is a basic principle in the theories of choice under investigation here, EU Theory and Portfolio Theory, and this algebraic property is equivalent in proba-

[3] EU Theory and Portfolio Theory make differential predictions only if the component gambles of a mixture are bilateral to the ideal level of risk. With no assumption about the risk order, laterality is unknown. This method of selecting components to form mixtures, then, is arbitrary and perhaps no better than random choice would be.

TABLE 2

Consistency and Intransitivity

| Subject | Consistency measure | Intransitive triples |
|---|---|---|
| 1 | 4.89 | 0 |
| 2 | 4.87 | 1 |
| 3 | 4.84 | 0 |
| 4 | 4.81 | 0 |
| 5 | 4.79 | 0 |
| 6 | 4.74 | 4 |
| 7 | 4.71 | 0 |
| 8 | 4.64 | 0 |
| 9 | 4.62 | 1 |
| 10 | 4.59 | 3 |
| 11 | 4.58 | 3 |
| 12 | 4.55 | 1 |
| 13 | 4.55 | 0 |
| Subtotal | 4.70 | 13 |
| 14 | 4.54 | 0 |
| 15 | 4.51 | 2 |
| 16 | 4.38 | 1 |
| 17 | 4.33 | 14 |
| 18 | 4.33 | 1 |
| 19 | 4.30 | 0 |
| 20 | 4.27 | 11 |
| 21 | 4.24 | 8 |
| 22 | 4.21 | 6 |
| 23 | 4.21 | 5 |
| 24 | 4.20 | 5 |
| 25 | 4.16 | 10 |
| 26 | 4.16 | 11 |
| Subtotal | 4.29 | 74 |
| Total | 4.49 | 87 |

bilistic terms to weak stochastic transitivity (Tversky, 1967). The total number of intransitive triples for each subject is reported in the second column of Table 2. For a diabolically intransitive subject, the maximum number of intransitive triples over the three sets of nine stimuli is 30 and the number expected by chance is 22. The maximum number for our subjects is 14 and the total over the 13 most consistent subjects is 13 and over the 13 least consistent subjects is 74. The amount of intransitivity is insignificant and its relation to inconsistency supports the conclusion that such intransitivity as occurs is a consequence of inconsistency.

*Test of EU Theory*

For each subject, the orderings of the 20 triples were classified as monotone, strictly folded, and inverted as defined above, and these three frequencies for each subject are reported in Table 3. The table is summarized by the two rows labelled Subtotal and the final row labelled Total.

TABLE 3
Classification of Response Patterns by $S$'s for Experiment I

| Subject | Satisfy EU theory & portfolio theory | Satisfy portfolio theory violate EU theory | Violate both theories |
|---|---|---|---|
| | monotone orderings | Strictly folded orderings | Inverted orderings |
| 1 | 17 | 1 | 2 |
| 2 | 16 | 0 | 4 |
| 3 | 15 | 2 | 3 |
| 4 | 13 | 4 | 3 |
| 5 | 10 | 2 | 8 |
| 6 | 12 | 6 | 2 |
| 7 | 13 | 5 | 2 |
| 8 | 13 | 4 | 3 |
| 9 | 2 | 10 | 8 |
| 10 | 5 | 13 | 2 |
| 11 | 14 | 0 | 6 |
| 12 | 5 | 14 | 1 |
| 13 | 7 | 12 | 1 |
| Subtotal | 142 | 73 | 45 |
| 14 | 8 | 9 | 3 |
| 15 | 9 | 5 | 6 |
| 16 | 12 | 4 | 4 |
| 17 | 12 | 8 | 0 |
| 18 | 7 | 6 | 7 |
| 19 | 12 | 6 | 2 |
| 20 | 11 | 8 | 1 |
| 21 | 10 | 7 | 3 |
| 22 | 13 | 2 | 5 |
| 23 | 11 | 4 | 5 |
| 24 | 9 | 5 | 6 |
| 25 | 17 | 2 | 1 |
| 26 | 10 | 0 | 10 |
| Subtotal | 141 | 66 | 53 |
| Total | 283 | 139 | 98 |

Using the frequency, 98, of the inverted orderings as an estimate of the error rate, we find the frequency of strictly folded orderings, 139, significantly in excess, ($\chi_1^2 = 7.08$, $p < 0.01$). So the violations of EU Theory may be partitioned into one set (inverted orderings) which are unaccounted for except as random error, and another, significantly larger, set (strictly folded) which have a predicted regularity compatible with Portfolio Theory. The conclusion is further substantiated by the fact that it is more firmly supported by the more consistent subjects and is diluted but still true of the less consistent subjects.

An analysis by subjects reveals that the excess of violations of EU Theory of the particular kind that is compatible with Portfolio Theory, is not uniformly distributed over subjects. Most of the disparity between strictly folded orderings and random violations of EU Theory is contributed by some six *S*s (#10, 12, 13, 14, 17, 20) as may be seen in Table 3. These *S*s have a total of 64 strictly folded orderings and only 8 random. This unequal distribution over *S*s is to be expected because without a prior knowledge of the risk order and a Ss' optimum level of risk there is no way to insure that the mixture for a particular subject will be constructed from gambles which bracket his optimum level of risk. In other words this is not a small effect accumulated over a lot of Ss but rather one that is of substantial magnitude in the context of alternatives appropriate to each *S*.

## 3. Experiment II

*Subjects*

The subjects were 50 student volunteers paid for participation in the experiment.

*Stimuli*

The gambles used in this experiment were generated from a basic set of four gambles all of which had the same three outcomes, —\$5, 0, \$5, but differed in their probabilities for the outcomes as indicated in Table 4 where gamble *A*, for example, is shown to have a probability of 0.9 that the outcome would be zero otherwise to win or lose \$5 was equally likely. It is a reasonable presumption that these gambles are ordered in risk in order from *A* to *D*.

Note that gambles *B* and *C* can each be obtained as a contingent mixture in two different ways. The simplest example is *C* as a 50/50 mixture of *B* and *D* ($C = B\ 1/2\ D$), i.e., to play *C* once is identically equivalent in terms of final outcomes and their respective probabilities to tossing a fair coin to play either *B* or *D* and then playing *B* or *D* as the coin indicates. More completely, it is easily seen that

$$B = A\ 2/5\ C = A\ 4/7\ D$$
$$C = B\ 1/2\ D = A\ 2/7\ D.$$

TABLE 4

The Basic Set of Gambles for Experiment II

| | Outcomes | | |
| --- | --- | --- | --- |
| | −\$5 | 0 | \$5 |
| *D* | 0.4 | 0.2 | 0.4 |
| *C* | 0.3 | 0.4 | 0.3 |
| *B* | 0.2 | 0.6 | 0.2 |
| *A* | 0.05 | 0.9 | 0.05 |

This set of four gambles was then used to generate four other sets by translating this basic set to expected value levels of \$1, \$2, \$3, and \$4 by adding the corresponding constant to all outcomes. For example the gamble *C* at expected value \$3 has outcomes −\$2, \$3, \$8 with probabilities 0.3, 0.4, 0.3 respectively, as before.

So at each of five expected value levels, 0 to \$4, we have a set of four gambles, and the members of each set all have the same outcomes and differ only in the probability distribution.

*Procedure*

The gambles in each set were presented pairwise on a slave scope. Each gamble was a circle with pie cuts to indicate the probabilities, and the outcomes indicated outside the circle next to its proper segment. Under the instructions that if he could play just one gamble once which would he choose, the subject indicated his preference by pushing a right or left button and then the computer erased the image. Then a sign saying "Push Black Button When Ready" appeared. When pushed it caused a new pair of gambles to appear.

Each subject was run for one hour periods on three successive days. The first day began with an instructional and practice session including playing the spinner program on the computer and then he went through one replication of the 30 paired comparisons, six for each of the five expected value levels, in a mixed order. On each of the second and third days the subject went through two replications of the 30 paired comparisons in a different order and with a break between replications, providing a total of five replications on each pairwise choice.

After each replication one game was played according to his choice on a pair. The pair of games was chosen with expected value conditional on whether he was ahead or behind from previous plays and to avoid excessive wins or losses. Once the pair of games was chosen by the experimenter it was displayed on the screen with a checkmark under the game the subject had indicated he preferred in that pair. It was then played on a homemade spinner wheel. The subject was paid at the end of the experiment an hourly wage of \$2.00 plus his winnings and losses over the five replications.

## Results

### *Measure of Consistency*

A measure of consistency was constructed for each subject as an average over the 30 pars of the 5–0, 4–1, or 3–2 splits exactly as in Experiment 1. The results were reported in Table 5 with the *S*s numbered and ordered from most to least consistent. The 0.05 level of significant consistency is 3.65 and the 0.01 level is 3.74. With an additional number of replications there is a potential problem about detecting true indifference. An even number of replications would, correspondingly, create a potential problem by obscuring a true preference. The level of consistency revealed in Table 5 suggests that the former is the lesser evil. Three additional *S*s with consistency measures of 3.60, 3.43, and 3.13 respectively were dropped on the grounds that their choices were not distinguishable from random behavior. The asterisks in Table 5 indicate *S*s who had monotone preference orderings everywhere as discussed below.

### *Construction and Classification of Preference Orderings*

The stochastically dominant choice on each pair if is transitive over the six pairs at each expected value level yields a rank ordered I scale. There were 22 of the 50 *S*s for whom the preference order was transitive at all five expected value levels and the preference order was monotone everywhere, i.e., every I scale was either *ABCD* or *DCBA*. These 22 *S*s who were monotone everywhere satisfy EU Theory perfectly in so far as the gambles used in this experiment are concerned. If the gambles are presumed to increase in risk from *A* to *D* then 15 of the 22 were risk aversive at all expected value levels, four were risk loving and three were risk averse at lower expected value levels and then changed to the other extreme at one of the higher expected value levels and stayed there.

The 28 *S*s who were not monotone everywhere would, of course, have one or more triples with a nonmonotone preference pattern and the classification of these triples is of particular interest. In Experiment I, the subject ordered each triple directly so the variety of orderings for the triple *A*, *B*, *C* for example, were classified as monotone (*ABC*, *CBA*), strictly folded (*BAC*, *BCA*), and inverted (*ACB*, *CAB*), for the reasons previously given. In Experiment II, in which the ordering on a triple is constructed from the dominant pairwise choices, a subject could yield, in addition to these three kinds of transitive orderings, either of two intransitive patterns, $A \to B \to C \to A$, $C \to B \to A \to C$.

So for every triple there are two possible response patterns for each of four kinds: monotone, strictly folded, inverted, and intransitive. Only the first, monotone, is compatible with EU Theory, the other three are all violations. If EU Theory is valid then the three kinds of violations should occur with equal frequency. To test this assumption for strictly folded and inverted orderings we note that if the true ordering

TABLE 5

Consistency Measures

| S# | Consistency measure | S# | Consistency measure |
|---|---|---|---|
| 1* | 5.00 | 26 | 4.46 |
| 2* | 4.96 | 27* | 4.33 |
| 3* | 4.96 | 28 | 4.33 |
| 4* | 4.96 | 29* | 4.33 |
| 5* | 4.93 | 30 | 4.30 |
| 6* | 4.93 | 31 | 4.26 |
| 7* | 4.90 | 32 | 4.26 |
| 8* | 4.90 | 33 | 4.23 |
| 9* | 4.87 | 34 | 4.23 |
| 10* | 4.87 | 35* | 4.17 |
| 11* | 4.83 | 36 | 4.17 |
| 12* | 4.83 | 37 | 4.10 |
| 13* | 4.80 | 38 | 4.10 |
| 14 | 4.77 | 39 | 4.06 |
| 15* | 4.73 | 40 | 4.06 |
| 16* | 4.73 | 41 | 4.03 |
| 17* | 4.67 | 42 | 4.00 |
| 18* | 4.67 | 43 | 3.97 |
| 19 | 4.63 | 44 | 3.97 |
| 20 | 4.63 | 45 | 3.97 |
| 21* | 4.57 | 46 | 3.93 |
| 22 | 4.57 | 47 | 3.93 |
| 23 | 4.57 | 48* | 3.93 |
| 24 | 4.53 | 49 | 3.93 |
| 25 | 4.46 | 50 | 3.73 |

* Indicates those *S*s who had monotone preference orderings at all five expected value levels.

ordering is *ABC*, then, if an error is made on the embedded pair *AB*, the resulting order will be *BAC* and will be strictly folded. If an error is made on the embedded pair *BC*, the resulting order will be *ACB* and will be classified as inverted. To test the assumption that these two kinds of errors are equally likely, we examined the inconsistency of the 22 monotone *S*s on all such pairs in every monotone triple, of which there were 440. The mean consistency level for 5 replications on the first embedded pair, such as *AB* in the triple *ABC*, is 4.680 and the mean consistency level on the second embedded pair, which is *BC* in that triple, was 4.718, not a significant difference.

Under Portfolio Theory the strictly folded orderings are also admissible and with a suitably designed experiment may be encouraged, so only the inverted and intransitive response patterns should still occur with equal frequency but significantly less than those which are strictly folded.

As a further control these 28 *S*s are divided into the 14 most consistent and the 14 least consistent. The results for each level of expected value are presented in Table 6 with the crucial subtotals at the bottom of the table. We find, overall, that there are 85 strictly folded orderings, violating EU Theory, which is over three times as many as expected on the basis of random error. We also see that the number of transitive orderings classified as inverted is the same as the number of intransitive patterns, as expected, and supports the contention that they represent random error.

TABLE 6

Classification of Response Patterns for Experiment II

| | | | Satisfy EU Theory and Portfolio Theory | Satisfy Portfolio Theory. Violate EU Theory | Violate both Theories | | |
|---|---|---|---|---|---|---|---|
| | | | Monotone Orderings | Folded Orderings | Inverted Patterns | | |
| | | | | | Transitive Orderings | Intransitive Patterns | |
| 22 Monotone *Ss* | | E=0 | 88 | 0 | 0 | 0 | |
| | | E=1 | 88 | 0 | 0 | 0 | |
| | | E=2 | 88 | 0 | 0 | 0 | |
| | | E=3 | 88 | 0 | 0 | 0 | |
| | | E=4 | 88 | 0 | 0 | 0 | |
| 28 Nonmonotone *Ss* | 14 Most Consistent | E=0 | 36 | 12 | 3 | 5 | |
| | | E=1 | 54 | 2 | 0 | 0 | |
| | | E=2 | 50 | 5 | 0 | 1 | |
| | | E=3 | 46 | 8 | 0 | 2 | |
| | | E=4 | 35 | 12 | 8 | 1 | |
| | 14 Least Consistent | E=0 | 44 | 7 | 2 | 3 | |
| | | E=1 | 43 | 11 | 1 | 1 | |
| | | E=2 | 35 | 14 | 4 | 3 | |
| | | E=3 | 39 | 8 | 4 | 5 | |
| | | E=4 | 40 | 6 | 4 | 6 | |
| | Subtotals | Most Cons. | 221 | 39 | 11 | 9 | |
| | | Least Cons. | 201 | 46 | 15 | 18 | |
| Total | | | 862 | 85 | 26 | 27 | 1000 |

## 4. Discussion and Interpretation

The absolute number of violations reported in Table 6 are not in themselves significant because they are dependent upon the particular *S*s selected, the range of gambles used at a fixed *E* level, the range of *E* levels used, and the particular probability mixtures used in constructing triples. This experiment, however, is representative in these respects of many of the experiments that have been reported studying risky decision making and is suggestive of one of the reasons for the viability of EU Theory. Overall 862 out of 1,000 response patterns on triples of gambles, 86.2 %, are compatible with EU Theory and there are not many behavioral theories that do as well. Portfolio Theory would raise this another 8.5 % to 94.7 %, and the rest, 5.3 %, is random error, as far as this experiment is concerned. The point that needs to be made is that if a theory is wrong in principle it will inevitably be wrong in some application no matter how useful it may be as an approximation pragmatically.

If expected utility theory is rejected as descriptive theory it means that no theory or procedure for measuring utility may exist which will justify maximizing expected utility as a theory of individual decision making under risk. A more complicated theory would have to be invented in which EU Theory was supplemented with something to account for experimental data, such as, for example, a risk preference. Tversky (1967) provided evidence that in comparing risky and riskless choices, a risk premium was needed to account for the risky choices with expected utilities based on the riskless choices, which, incidentally, was a *positive* increment, but his subjects were inmates of a prison.

The results indicate that the available procedures for measuring utilities (e.g., von Neumann & Morganstern, 1953; Savage, 1954; Davidson, Suppes, & Siegel, 1957; or Coombs & Komorita, 1958) cannot be safely employed since their assumptions are not satisfied by the subjects. There are systematic violations of EU Theory which can be accounted for by a more general theory. Expected utility theory can, of course, be used normatively to guide the individual in his choices.

## References

Adams, E. W., & Fagot, R. F. A model of riskless choice. *Behavioral Science*, 1959, **4**, 1–10.

Coombs, C. H. Portfolio theory: a theory of risky decision making. *La Decision*, Paris, France, Centre National de la Recherche Scientifique, 1969.

Coombs, C. H., & Bowen, J. Additivity of risk in portfolios. *Perception and Psychophysics*, 1971, **10**, 43–46.

Coombs, C. H., & Huang, L. C. Polynomial psychophysics of risk. *Journal of Mathematical Psychology*, 1970(a), **7**, 317–338.

Coombs, C. H., & Huang, L. C. Tests of a portfolio theory of risk preference. *Journal of Experimental Psychology*, 1970(b), **85**, 23–29.

COOMBS, C. H., & KOMORITA, S. S. Measuring utility of money through decisions. *American Journal of Psychology*, 1958, **71**, 383–389.

DAVIDSON, D., SUPPES, P., & SIEGEL, S. *Decision making: An experimental approach.* Stanford, California: Stanford University, 1957.

DUNN, J. G. Subjective and objective risk distribution, a comparison and its implication for accident prevention. *Occupational Psychology*, 1972, **46**, 183–187.

HUANG, L. C. The expected risk function. Michigan Mathematical Psychology Program technical report (MMPP) 1971-6, The University of Michigan, 1971(a).

HUANG, L. C. Experiments on the measurement of risk. Michigan Mathematical Psychology Program technical report (MMPP) 1971-7, The University of Michigan, 1971(b).

LEE, W. *Decision theory and human behavior.* New York: John Wiley and Sons, 1971.

LUCE, R. D., & RAIFFA, H. *Games and decisions.* New York: John Wiley and Sons, 1957.

POLLATSEK, A., & TVERSKY, A. A theory of risk. *Journal of Mathematical Psychology*, 1970, **7**, 540–553.

SAVAGE, L. J. *The foundations of statistics.* New York: John Wiley and Sons, 1954.

STARR, C. Social benefit versus technological risk. *Science*, 1969, **165**, 1232–1238.

TVERSKY, A. Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology*, 1967, **75**, 27–36.

VON NEUMANN, J., & MORGENSTERN, O. *Theory of games and economic behavior.* Princeton, New Jersey: Princeton University Press, 1st Ed., 1944; 3rd Ed., 1953.