

Testing Expectation Theories of Decision Making without Measuring Utility or Subjective Probability¹

C. H. COOMBS, T. G. BEZEMBINDER,² AND F. M. GOODE

The University of Michigan, Ann Arbor, Michigan

Certain empirical implications of those decision-making theories which involve maximizing an expectation are derived. These implications are all measurement-free, so neither subjective probability nor utility need be measured. Two experiments are reported, the first exploratory and the second intensive and substantial. EV, EU, and SEV theory were inadequate to account for the behavior of 12% or more of the subjects in either experiment. SEU theory was inadequate in 10% or less of the cases in the first experiment and 5% of the cases in the second experiment.

1. INTRODUCTION

There are four principal theories claiming that, in risky situations in which a given set of actions is available, people make decisions by choosing the act that offers the highest expectation. These four theories differ in the assumption that utility for money is linear or nonlinear with the objective value of money and that subjective probability is linear or nonlinear with objective probability. One of these theories, expected value or EV theory, is widely known: it defines "expectation" in terms of objective probability and objective value of money. A theory defining "expectation" in terms of linear transformations of these components, such as linear subjective probability and linear utility for money, would coincide with EV theory. The other three combinations are known as (1) expected utility or EU theory, which defines expectation in terms of objective probability and nonlinear utility for money; (2) subjectively expected value or SEV theory, involving nonlinear subjective probability and linear utility for money; and, finally; (3) subjectively expected utility or SEU theory, involving both nonlinear subjective probability and nonlinear utility for money (Edwards, 1954).

The early literature of decision making has been reviewed by Edwards in 1954. An extended review of the literature covering the period from 1954 to 1960 has been made available by the same author in 1961. Consequently we will only attempt to summarize the present status.

¹ This research was supported by Public Health Service Research grant MH-04236.

² Present address: University of Nijmegen, The Netherlands.

Quite in contrast to its strength as a *normative* theory, maximizing expected value has never been taken seriously as a *descriptive* theory of decision-making under risk. Ever since Daniel Bernoulli in 1738 introduced the concept of nonlinear utility for money ("moral value", in his terms) in his discussion of the St. Petersburg paradox (Bernoulli, 1954; see also Chernoff and Moses, 1959), other theories have predominated.

The other three theories have been given some substantial amount of experimental attention. Some of the principal experimental studies of one or more of these models include the following: Coombs and Komorita (1958), Davidson and Marschak (1959), Davidson, Suppes, and Siegel (1957), Edwards (1955), Hurst and Siegel (1956), and Mosteller and Nögee (1951). In these studies one typically uses an individual's preferential choices among gambles to construct a functional curve for utility and/or subjective probability—which of these curves is constructed depends on the model under scrutiny—and, assuming the model holds, these curves are used to predict his choices among some other gambles. Generally the predictions are substantially confirmed and the SEU model appears somewhat superior to any of the others. It should be noted, however, that by assuming nonlinearity in both components of the gamble the SEU model introduces more degrees of freedom and therefore cannot possibly fit the data worse than any of the other models that put more constraints on the data.

Every one of these experiments is subject to severe criticism on one ground or another (see Edwards, 1961), and the overall picture is inconclusive. Part of the difficulty lies in the tremendous generality of the SEU model. Every subject is a separate experiment, as his utility function and subjective probabilities may be different from those of others; hence, considerable experimental effort usually has to be expended pinning down these values and this may have undesirable consequences, e.g., boredom on the part of the subjects. Other potential bases for criticism of such experiments include the following: departure from reality in decision making if gambles are hypothetical, confounding effects of changes in the status quo if gambles are not hypothetical, changes in utility or subjective probability if the experiment runs over an extended period, etc.

One possible way out is, at the end of each experimental session, to select some gambles reflecting the subject's individual preferences and genuinely play them. If an adequate interval between experimental sessions is allowed, effects of changes in status quo from one session to the next may be neglected. This was done in the second of the two experiments reported below.

The section that follows presents some simple theorems which permit testing these several models in a single experiment without measuring utilities and/or subjective probabilities. The section concludes with a discussion of the implications of the theorems for the experimental design and the analysis of the resulting data matrix. The third section presents Experiment I and a detailed discussion of the theory underlying the statistical analysis. The last section presents Experiment II.

2. THEOREMS

Notation:

p_j = an objective probability;

ψ_j = a subjective probability, a function of p_j ;

v_j = an amount of money;

u_j = the utility of an amount of money v_j ;

G_j = a gamble with two outcomes: $v_j (> 0)$ and zero (i.e., no money); the gamble G_j offers a probability p_j to win v_j and a probability $1 - p_j$ to win zero; therefore, we will write $G_j = (p_j, v_j)$.

The expectation, E_j , for such a gamble under each of the four theories is as follows.

Expected value theory:

$$E_j = p_j v_j ; \quad (1)$$

Expected utility theory:

$$E_j = p_j u_j ; \quad (2)$$

Subjectively expected value theory:

$$E_j = \psi_j v_j ; \quad (3)$$

Subjectively expected utility theory:

$$E_j = \psi_j u_j . \quad (4)$$

In this study we will consider pairs of gambles of the form G_j . These pairs have the characteristic that one of the gambles of the pair offers a higher probability to win a smaller payoff than the other. The former gamble, that is the one with the higher probability of winning the smaller amount, will be designated by the subscript l , for left, and the latter by the subscript r , for right.³ Hence, we have

$$G_l(p_l, v_l) \text{ and } G_r(p_r, v_r) \text{ where } p_l > p_r \text{ and } v_l < v_r .$$

In such a pair of gambles we are particularly interested in the ratio of the expectations of the gambles. We will study how these ratios change as we add a constant positive probability increment, $\Delta p (> 0)$, to both p_l and p_r and/or a constant positive increment in the objective value of the money payoff, $\Delta v (> 0)$, to both v_l and v_r . The components of the gambles to which increments are to be added (probability of winning and/or money value of the payoff) will depend upon the model under consideration.

³ See Fig. 1 for the motivation for labeling. In this figure the left and right gambles that form a pair are connected by a dotted line.

We will let a prime attached to a quantity always indicate the addition of an increment to the corresponding unprimed quantity in a pair of gambles. Thus

$$p_l' - p_l = \Delta p = p_r' - p_r,$$

and

$$v_r' - v_r = \Delta v = v_l' - v_l.$$

We define four pairs of gambles, a, b, c, d , as follows (cf. Figure 1):

$$\begin{aligned} a &= [G_l(p_l, v_l), G_r(p_r, v_r)], \\ b &= [G_l(p_l', v_l), G_r(p_r', v_r)], \\ c &= [G_l(p_l, v_l'), G_r(p_r, v_r')], \\ d &= [G_l(p_l', v_l'), G_r(p_r', v_r')]. \end{aligned} \tag{5}$$

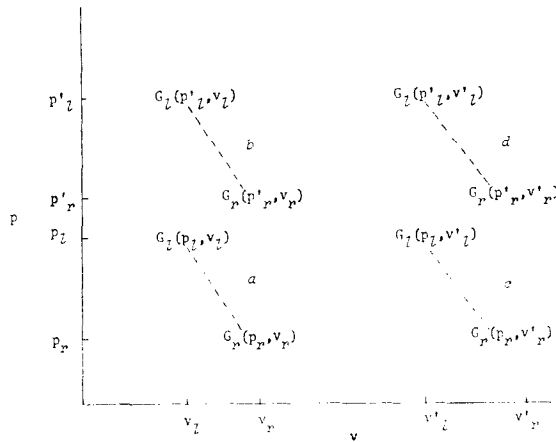


FIG. 1. Illustration of four pairs of gambles.

We shall assume that subjective probability, ψ , is strictly monotonic with objective probability and that utility for money is strictly monotonic with money. Hence

$$u_j' = u(v_j') = u(v_j + \Delta v) > u_j = u(v_j),$$

and

$$\psi_j' = \psi(p_j') = \psi(p_j + \Delta p) > \psi_j = \psi(p_j).$$

Similarly,

$$u_l < u_r \quad \text{if and only if} \quad v_l < v_r,$$

and

$$\psi_l > \psi_r \quad \text{if and only if} \quad p_l > p_r.$$

For each pair of gambles we may form the ratio of their expectations as defined by any one of Eqs. 1-4. Under subjectively expected utility, for example, and the pair of gambles a , we may write

$$\psi_i u_i / \psi_r u_r = A. \quad (6)$$

As another example we may take the ratio C for the pair of gambles c ; under expected utility theory:

$$p_i u_i' / p_r u_r' = C.$$

In this way we will have for each theory the ratios A , B , C , and D for the corresponding pairs of gambles a , b , c , and d . It will not be necessary to have a distinguishing notation for these ratios under the different expectation theories as they will never occur in conjunction. The theorems to be proven pertain to the relations among these ratios for the various pairs of gambles.

In summary, the gambles considered here are two outcome gambles with one outcome zero. Pairs of such gambles are formed in which the left gamble of the pair offers a higher probability of winning a smaller amount. We form the ratio of the expectations in such a pair of gambles and will next consider how the ratio changes if a new pair is formed by adding a constant positive increment to the probabilities of winning and/or a constant positive increment to the values to be won.

THEOREM 1. *For the pairs of gambles of the form considered, and with $v_i, v_r \neq 0$, $p_r > 0$, and $\Delta p > 0$, then EV theory requires that $A > B$; where*

$$p_i v_i = A p_r v_r,$$

and

$$p_i' v_i = B p_r' v_r.$$

Proof.

$$A = \frac{p_i}{p_r} \cdot \frac{v_i}{v_r},$$

$$B = \frac{p_i'}{p_r'} \cdot \frac{v_i}{v_r} = \frac{p_i + \Delta p}{p_r + \Delta p} \cdot \frac{v_i}{v_r},$$

so that

$$\frac{A}{B} = \frac{p_i(p_r + \Delta p)}{p_r(p_i + \Delta p)} > 1,$$

as

$$p_i \Delta p > p_r \Delta p.$$

Therefore

$$A > B.$$

Q.E.D.

Interpretation: If we have a pair of gambles oriented as the pair a in Fig. 1, in which neither v_i nor v_r is zero and $p_r > 0$, then the effect of adding a fixed increment in probability, $\Delta p > 0$, to both p_i and p_r is to construct another pair of gambles, b ,

in which the ratio of the expectation of the left gamble to the expectation of the right gamble is smaller than the corresponding ratio in the pair a .

Consider next the effect of adding a fixed positive increment Δv to the amount to be won in each gamble of a pair.

THEOREM 2. *For pairs of gambles of the form considered, and with $v_r v_l \neq 0$, $p_r > 0$, and $\Delta v > 0$, then EV theory requires that $C > A$; where*

$$p_l v_l = A p_r v_r,$$

and

$$p_l v_l' = C p_r v_r'.$$

Proof. If we take the ratio C/A , we find that $C > A$ by the same reasoning as was followed to prove Theorem 1.

Interpretation: If we have a pair of gambles oriented as the pair a in Fig. 1, in which neither v_l nor v_r is zero and $p_r > 0$, then the effect of adding a fixed increment in money value, $\Delta v > 0$, to both v_l and v_r is to construct another pair of gambles, c , in which the ratio of the expectation of the left gamble to the expectation of the right gamble is greater than the corresponding ratio in the pair a .

Immediately, from Theorem 1, we have, *mutatis mutandis*, Theorem 3 for EU theory; and from Theorem 2 we have Theorem 4 for SEV theory.

THEOREM 3. *For pairs of gambles of the form considered, and with $u_l u_r \neq 0$, $u(v = 0) = 0$, $p_r > 0$, and $\Delta p > 0$, EU theory requires that $A > B$; where*

$$p_l u_l = A p_r u_r,$$

and

$$p_l' u_l = B p_r' u_r.$$

THEOREM 4. *For gambles of the form considered, with $v_r v_l \neq 0$ and $\psi(p > 0) > 0$, and $\Delta v > 0$, SEV theory requires that $C > A$; where*

$$\psi_l v_l = A \psi_r v_r,$$

and

$$\psi_l v_l' = C \psi_r v_r'.$$

Finally we have Theorem 5 for SEU theory.

THEOREM 5. *For gambles of the form considered, and if $u(v \neq 0) \neq 0$ and $\psi(p > 0) > 0$, SEU theory requires that $BC = AD$; where*

$$\psi_l u_l = A \psi_r u_r,$$

$$\psi_l' u_l = B \psi_r' u_r,$$

$$\psi_l u_l' = C \psi_r u_r',$$

and

$$\psi_i' u_i' = D\psi_r' u_r'.$$

Proof.

$$\frac{(\psi_i' u_i)(\psi_i u_i')}{(\psi_i u_i)(\psi_i' u_i')} = \frac{BC(\psi_r' u_r)(\psi_r u_r')}{AD(\psi_r u_r)(\psi_r' u_r')},$$

so that

$$\frac{BC}{AD} = \frac{\psi_i' u_i \psi_i u_i' \psi_r u_r \psi_r' u_r'}{\psi_i u_i \psi_i' u_i' \psi_r' u_r \psi_r u_r'} = 1. \quad \text{Q.E.D.}$$

We will assume that the individual's preference on any pair of gambles will indicate whether the ratio of the appropriate expectations is greater or less than one. Hence, we make the following correspondence to experimental data on the pairwise preferences under any one of the theories assuming an individual maximizes the appropriate expectation:

$$\begin{aligned} G_i(p_i, v_i) \text{ is preferred to } G_r(p_r, v_r) &\Leftrightarrow A > 1, \\ G_i(p_i', v_i) \text{ is preferred to } G_r(p_r', v_r) &\Leftrightarrow B > 1, \\ G_i(p_i, v_i') \text{ is preferred to } G_r(p_r, v_r') &\Leftrightarrow C > 1, \\ G_i(p_i', v_i') \text{ is preferred to } G_r(p_r', v_r') &\Leftrightarrow D > 1. \end{aligned} \quad (7)$$

The method of analysis to be discussed in a later section is somewhat complicated and it will be useful here to first illustrate the experimental implications of these theorems for a small hypothetical data matrix. If an individual chooses among the gambles of a pair and makes his choice according to any expectation maximization theory, then he should prefer the left gamble more strongly to the right one as the ratio of the expectation of the left gamble to the expectation of the right one increases. In accord with EU theory, he should, for example, prefer the left gamble more strongly to the right gamble in the pair *a* than in the pair *b* (Theorem 3). He certainly satisfies this requirement, if preferring the left gamble to the right one in the pair *a*, he makes the opposite preferential choice in the pair *b*. Obviously, this also applies to, respectively, the pairs *c* and *d*. We will call this behavior a confirmation of EU theory.

Designating a preferential choice of a left gamble over a right gamble by the symbol 1 and a choice of a right over a left gamble by the symbol 0, and substituting these symbols for the pairs of gambles in Fig. 1, we would have the following little matrix of these hypothetical data:

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

The first column in this data matrix represents a data pattern that confirms EU theory for the pairs *a* and *b* in Fig. 1. Similarly, the second column shows a confirming pattern for the pairs *c* and *d*. The other possible EU patterns that may appear in a

column, are $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The pattern "1 above a 0", i.e., $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, is a violation of EU theory because if an individual chooses the right gamble over the left in the pair a (designated by 0) and makes the opposite choice in the pair b , (designated by 1), then he more strongly prefers the left gamble over the right one in the pair b than in the pair a which is the opposite of what is required by EU theory according to Theorem 3.

It is easy to see that the rows of the above little data matrix may be used to test SEV theory. Because SEV theory requires $C > A$ (Theorem 4), an individual's preference for the left gamble to the right one should be stronger in the pair c than it would in the pair a . The row pattern $\begin{bmatrix} 0 & 1 \end{bmatrix}$ is, therefore, a confirmation of SEV theory; the pattern $\begin{bmatrix} 1 & 0 \end{bmatrix}$ is a violation.

As the patterns $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \end{bmatrix}$ do neither confirm nor violate, respectively, EU and SEV theory, they will be called *compatible* patterns.

Now by changing both components, i.e., probabilities and payoffs, of the gambles in the pair a at once, we have the pair d . The test for SEU theory involves the entire 2×2 matrix (or any second-order minor of a larger data matrix) in accordance with Theorem 5. As each of the four cells of this matrix may show a 1 or a 0, there are $2^4 = 16$ possible SEU patterns. Which of these patterns are confirming, violating or compatible with SEU theory, will be discussed in the section on method of analysis.

It is not very difficult to expand our little basic data matrix. If we take a pair of gambles such as a in Fig. 1, we can construct a set of pairs of gambles by adding successive increments Δp to both p_l and p_r , as illustrated in Fig. 2. In this way we can generate a column of a data matrix. By adding, in an analogous manner, successive increments Δv to both v_l and v_r of the gambles in the pair a , we may construct a row of a data matrix. Constructing in this way an $r \times k$ data matrix, we will have in each of its k columns at most $\binom{r}{2}$ tests for EU theory, and in a column we will get this maximum number of tests, as a consequence of Theorems 1 and 3, only if the change Δp is the same fixed increment added to p_l and p_r in a pair of gambles. In each of the r rows of the $r \times k$ data matrix we will similarly have at most $\binom{k}{2}$ tests for SEV theory, and we will only get all of these if Δv is the same increment added to v_l and v_r in a pair of gambles (cf. Theorems 2 and 4). Finally, the $r \times k$ data matrix contains $\binom{r}{2}\binom{k}{2}$ second-order minors, and each of these provides us with a test of SEU theory, as Theorem 5 does not require Δp or Δv to be the same constant in, respectively, pairs of rows or pairs of columns. It should be noted, however, that the tests for EU, SEV, and SEU theory are in general not independent.

It goes without saying that each cell of our $r \times k$ data matrix provides us with a statistically independent test for EV theory (Eq. 7). In constructing our experimental set of gambles we may make an appropriate choice of p_l and p_r and of the numerical payoffs v_l and v_r such that the ratio of the *expected value* of the left gamble to that of the right gamble is greater than unity ($A > 1$ in Theorem 1) for the pair with the lowest probabilities; and this ratio can be made to decrease with successive increments

Δp to a value less than unity ($B < 1$ in Theorem 1). As EV theory predicts what choice the subject should make in any one pair and therefore what symbol, 1 or 0, should appear in any one cell of the data matrix, it predicts where going up in a column the symbols should change from 1 to 0. EV theory also predicts where, going to the right in a row, the symbols should change from 0 to 1. Because the function relating utility to money is unknown, EU theory does not predict the actual position of the cutoff point between 1's and 0's in a column of the data matrix. Analogously, in the absence of the function relating subjective to objective probability, SEV theory will not predict the actual cutoff point between 1's and 0's in a row.

3. EXPERIMENT I

A. EXPERIMENTAL DESIGN

Stimuli

The stimuli were gambles of the form discussed in the previous section: they offered a probability, p_j , to win a finite amount of money v_j —otherwise nothing. The particular gambles used and their organization into pairs and sets of pairs are perhaps most immediately clear from Fig. 2. There are four sets of gambles, each set consisting of eight pairs which differ systematically (see Table 1 also). In Set I, for example, v_l is \$0.80 and v_r is \$1.20; $p_l - p_r = 0.1$; and between successive pairs

$$\Delta p = p_l' - p_l = p_r' - p_r = 0.1.$$

The other sets differ from Set I only in v_l and v_r ; in Set II we have $v_l = \$1.70$, $v_r = \$2.30$; Set III: $v_l = \$2.80$, $v_r = \$3.20$; Set IV: $v_l = \$3.70$, $v_r = \$4.30$.

The expected values of the gambles in Sets I–IV are given in Table 1. The arrangement of the cells in this table is similar to the arrangement of the gambles in Fig. 2. So, for example, the gambles comprising the top pair in Set I have the expected values given in the top cells of columns 1 and 2, i.e., \$0.72 and \$0.96. Note that the right gamble has the higher expected value. In the bottom pair of gambles in Set I, however, we have expected values \$0.16 and \$0.12 and so, here, the left gamble has the higher expected value (see Theorem 1).

From Table 1 it may also be seen that the increase Δv is only a constant going from Set I to III and from Set II to IV. In both pairs of sets, $\Delta v = \$2.00$, so only these pairs provide us with tests of SEV theory. This is the price paid for having a sensitive design for EU theory, which was constructed by putting the EV cutoff point low in Sets I and II and high in Sets III and IV. Hopefully, we hit the EU cutoff points of the subjects with positively and negatively accelerating utility functions, respectively, by the pairs of gambles in the Sets I and II, and in III and IV.

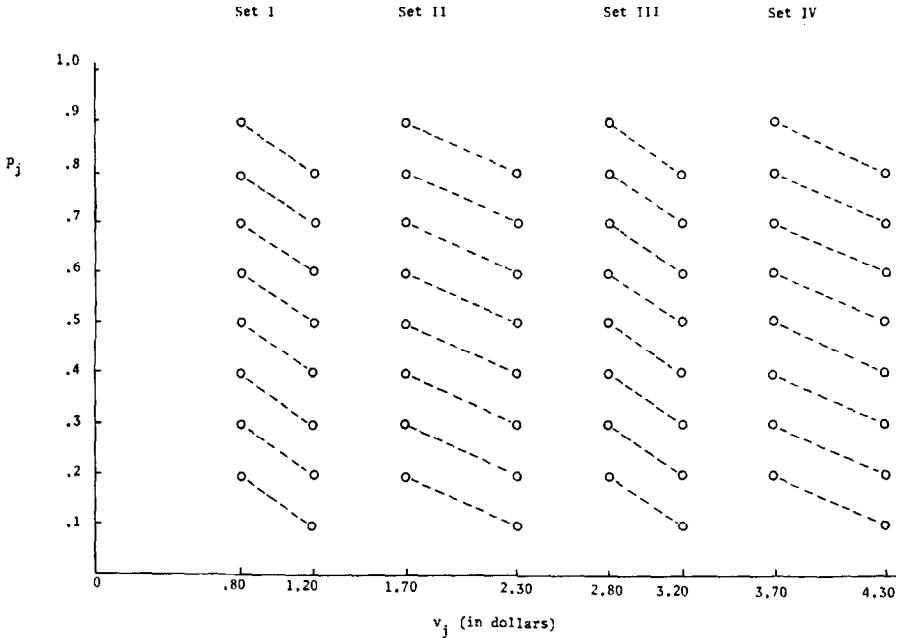


FIG. 2. The four sets of pairs of gambles in Experiment I.

TABLE 1
EXPECTED VALUES OF GAMBLES USED IN EXPERIMENT I

Prob. ^a	Amount of Money To Be Won ^b							
	Set 1		Set 2		Set 3		Set 4	
	0.80	1.20	1.70	2.30	2.80	3.20	3.70	4.30
0.9	0.72	—	1.53	—	2.52	—	3.33	—
0.8	0.64	0.96	1.36	1.84	2.24	2.56	2.96	3.44
0.7	0.56	0.84	1.19	1.61	1.96	2.24	2.59	3.01
0.6	0.48	0.72	1.02	1.38	1.68	1.92	2.22	2.58
0.5	0.40	0.60	0.85	1.15	1.40	1.60	1.85	2.15
0.4	0.32	0.48	0.68	0.92	1.12	1.28	1.48	1.72
0.3	0.24	0.36	0.51	0.69	0.84	0.96	1.11	1.29
0.2	0.16	0.24	0.34	0.46	0.56	0.64	0.74	0.86
0.1	—	0.12	—	0.23	—	0.32	—	0.43

^a Indicates probability of winning. — ^b In dollars and cents.

Subjects

Two distinct groups of subjects were run in Experiment I. Group 1 comprised 18 mature adults, all Negroes, living in Willow Village, near Ypsilanti, Michigan.

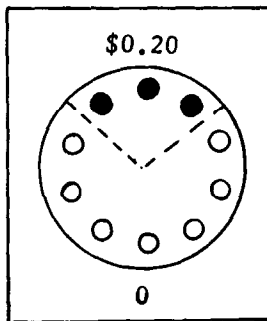
There were ten men and eight women, a number of whom were man and wife. They were asked by word of mouth if they would care to participate in a study of decision making. They were run as a group in one experimental session of $2\frac{1}{2}$ hours for which they were paid \$4.00 each.

Group 2 comprised 22 University of Michigan undergraduates who had previously listed themselves as available as paid subjects for psychological experiments. There were nine male students and 13 female students. They were also run as a group in one evening session of $2\frac{1}{2}$ hours.

Procedure

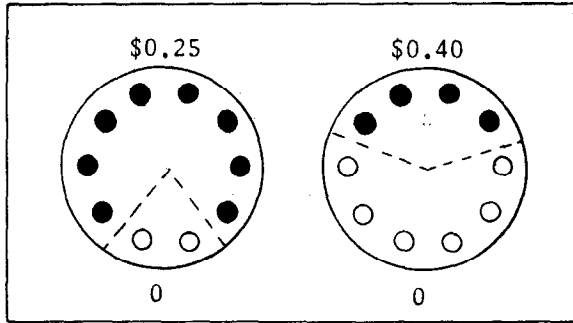
The two groups were run using identically the same procedure except that the senior author conducted the session for group 1 and the second author conducted the session for group 2. In the experimental session each subject went through six booklets. Three of the booklets (1, 3 and 5) contained the 32 pairs of gambles and a preferential choice was obtained on each pair. The other three booklets (2, 4 and 6) contained the 64 distinct gambles, presented singly, and the instructions were different for each booklet. In these three booklets the subject was asked what is the most he would pay for a ticket to play the gamble (booklet 2, "buy" data); what is the least for which he would sell a ticket to play the gamble (booklet 4, "sell" data); and finally, he was asked to set a price on a ticket not knowing whether he would have to buy or sell it at that price (booklet 6, "price" data).

Booklet 1. The instruction page for the first booklet was as shown below.



This is a special kind of gamble. It is kind of a wheel of fortune with 10 pegs. It has a spinner that goes around and stops on one of the pegs. If it stops on one of the 3 black pegs you get \$0.20; if it stops on one of the 7 blank pegs you get nothing. Now look at this pair of gambles.

The gamble on the left is a wheel of fortune that has 8 black pegs and 2 blank ones. You would win \$0.25 if the spinner stopped on any one of the 8 black pegs. The wheel of fortune on the right has 4 black pegs and if the spinner stopped on any one of those, you would win \$0.40.



Now, if you could play only one of these two gambles, which one would you choose? Put your checkmark inside the wheel you choose.

On the following pages there are several pairs of gambles such as these. There will be 2 wheels of fortune and they will always have their winning black pegs at the top with the amount that you would win if the spinner stopped on one of the black pegs.

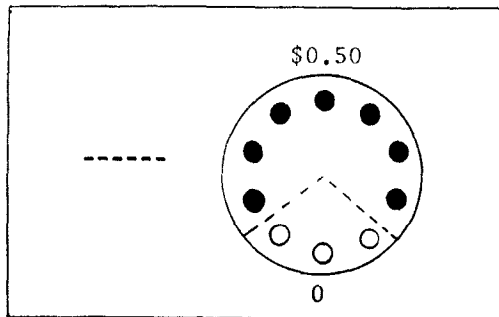
Please, look at each pair on the following pages, and decide which gamble you would rather play if you could play only one of them. Put the checkmark inside the wheel of the gamble you choose. There are no right or wrong answers; we just want to know your choices.

While you are thinking about one pair of gambles, you should forget any choice you made earlier, because all pairs are different.

Two wheels of fortune were used to illustrate the instructions and all questions were freely answered. The rest of the booklet consisted of four pages on each of which were eight pairs of gambles, each pair displayed in a box as illustrated in the instructions. The 32 pairs of gambles, of course, were those indicated in Fig. 2 and Table 1.

The 32 pairs of gambles were in a random order and the pairs were displayed in such a way as to balance any response bias for the left or right member of a display.

Booklet 2. The instruction page for the second booklet was as shown below.



This wheel of fortune has 7 black pegs and it gives you a good chance to win \$0.50. Now, imagine that you have to have a ticket to play this gamble. The ticket gives you one play on this wheel of fortune, and depending upon where the spinner stops you get \$0.50 or nothing. What is the highest price you would be willing to pay for a ticket for this gamble? Write that price on the dotted line to the left of the wheel.

On the next few pages you will find some more gambles of this kind. Please write down on the left of each wheel the highest price you would be willing to pay for a ticket for that gamble.

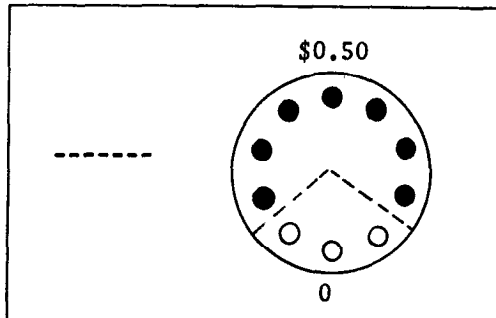
There are no right or wrong answers and you should only think about one gamble at a time.

To make clear that it was the *highest* price they would be willing to pay which was wanted, an assistant was asked by the experimenter what he would pay for an illustrative gamble (one not used in the experimental series), and then questioned whether he wouldn't pay more, until his ceiling was reached. Again, all questions were freely answered before proceeding with the experimental series.

The rest of the booklet consisted of eight pages on each of which were eight of the 64 distinct gambles. The sequence of 64 was arranged to balance the four sets of gambles.

Booklet 3. The instruction page for Booklet 3 was the same as that for Booklet 1 and the instructions were briefly reviewed. The four pages which followed contained the 32 pairs and were the same pages as those for Booklet 1 but permuted in the order 2, 3, 4, 1.

Booklet 4. The instruction page for Booklet 4 was as shown below.



Suppose you have a ticket to play this gamble. The ticket gives you only one play on this wheel of fortune. So, if you sell your ticket to someone else, you cannot play this gamble any more. What is the smallest amount for which you would be willing to sell your ticket to someone else? Write that amount on the dotted line to the left of the wheel.

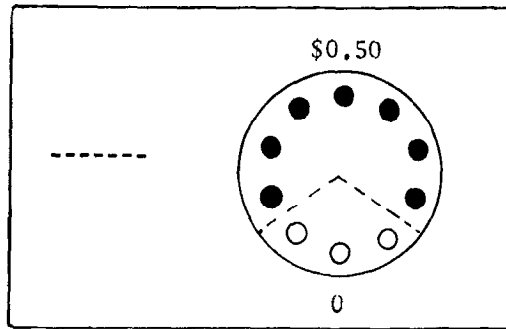
Please do the same thing with the gambles on the following pages. Again, there are no right or wrong answers and you should only think about one gamble at a time.

Again an assistant was used to make clear that they were asked to give the *lowest* price at which they would sell a ticket to play the gamble.

The following eight pages were the same as those of Booklet 2 but permuted in the order 5, 6, 7, 8, 1, 2, 3, 4.

Booklet 5. The instruction page was the same as that for Booklet 1 and again the instructions were briefly reviewed. The four pages which followed were the same as those for Booklet 1 but permuted in the order 3, 4, 1, 2.

Booklet 6. The instruction page for Booklet 6 was as shown below.



Suppose you have a ticket to play this gamble and your neighbor also has a ticket for it. Now, would you please put down one single amount that may be used either for buying your neighbor's ticket or selling your own. When you put down this amount, you don't know whether it is going to be used for buying or selling. Just pretend that we will decide whether the amount you have written down will be your buying price or your selling price. If we decide the amount is a buying price, then you would need to pay your neighbor that price for his ticket. But if we decide that your amount is a selling price, then you would have to sell your ticket for that price. Please write down, now, that amount on the dotted line on the left of the wheel.

On the following pages you will find some other gambles. For each gamble you should figure out a single price that may be used either for buying your neighbor's ticket or selling your own. For each gamble you should write down that amount without knowing which way it will be used.

The eight pages which followed contained the experimental series and were the same as those of Booklet 2 but permuted in the order 8, 7, 6, 5, 4, 3, 2, 1.

B. METHOD OF ANALYSIS

Pair-Comparison Data

On each pair of gambles three replications of the subject's preferential choice were obtained. For each subject these three choices were summarized in an 8×4

cumulative data matrix showing in each of its 32 cells—which represented the 32 pairs of gambles as displayed in Fig. 2—the number of times in the three replications that he preferred the left gamble over the right one, that is an integer from the set $\{0, 1, 2, 3\}$. By mapping frequencies of 0 and 1 into a 0 and frequencies of 2 and 3 into a 1 the cumulative matrix could be readily converted into a *reduced* (0, 1) matrix of the subject's stochastically dominant choices.

If a theory holds in an absolute and deterministic manner, then replications are unnecessary, and only the integers 0 and 3 would appear in the cumulative matrix. This was manifestly not the case, and the question arises whether a theory can be said to hold in a stochastic sense, in the sense of failing to hold only due to random fluctuations in judgment.

Because of the mutual dependencies between the different data patterns, it was not clear how the reduced matrix—or any data matrix—could be used to assess that question. Before going into details we may, therefore, first give, in the paragraph below, a rather general idea of the tests we devised for any particular subject using his cumulative and reduced matrices. As the entire following presentation refers to the i th subject we will omit the subscript i on the parameters to be used.

Corresponding to a manifest choice of the subject, we assume there exists a latent choice, which he would manifest in case there were no random fluctuations in judgment. Due to fluctuations of this kind, however, his manifest and latent choices may sometimes be different, and thus he may not be completely consistent over replications in his choices on a given pair of gambles. Since he makes binary choices, for a given pair of gambles at least one-half of his choices will be the same over replications, regardless of how much the choices agree with his latent preference. Thus for a given pair of gambles his consistency over replications must be between 0.5 and 1.0.

We will obtain, within this range, two independent estimates of the subject's consistency. The first estimate, \hat{p} , will be independent of any expectation theory. The other estimate, \tilde{p} , will depend upon the particular expectation theory to be tested and may, therefore, assume different values for different expectation theories. If $\tilde{p} < \hat{p}$, the expectation theory apparently does not put enough constraint on the data but rather requires a "fudge" factor of inconsistency on the subject's part that exceeds his level of inconsistency estimated independently of any expectation theory. So, if for a particular subject $\tilde{p} < \hat{p}$, the expectation theory taken into account by \tilde{p} is clearly to be rejected.

In order to estimate \hat{p} we first define π as the probability that on a given pair of gambles the subject's choice in any particular replication is the same as his true preference. This probability π is independent from any expectation theory and expressing the probability of a manifest choice to reflect the underlying true choice it may be regarded as a consistency index for any particular replication.

The π for an individual may be readily estimated from the following considerations. The cumulative data matrix of an individual has a 3 or a 0 in a cell if and only if the

individual makes the same choice on all three replications. The probability that he makes his true latent choice three times is π^3 ; the probability for him to make the opposite choice in all three replications is $(1 - \pi)^3$. If we let λ be the proportion of the entries that are 3 or 0 in his cumulative matrix, we may write the following equation:

$$\pi^3 + (1 - \pi)^3 = \lambda. \quad (8)$$

Similarly, the cumulative matrix shows an entry of 1 or 2 whenever the individual makes a particular choice in any one and the opposite choice in the two remaining replications. In either case his single odd choice can, of course, occur on any one of the three replications. As $1 - \lambda$ is the proportion of 1's and 2's we have the equation:

$$3\pi^2(1 - \pi) + 3\pi(1 - \pi)^2 = 1 - \lambda.$$

This equation may be reduced to Eq. 8 and has the same solution:

$$\pi = \frac{1}{2} \pm \frac{1}{6} \{[3(4\lambda - 1)]^{1/2}\}. \quad (9)$$

The consistency estimated by π must lie in the interval 0.50-1.00, where the lower bound represents consistency at the level of tossing a fair coin. We don't demand that these theories predict random behavior, but some significant level of consistency is required to make a fair test. For this reason, π values above 0.68 (beyond 2σ) will be taken to represent significant levels of consistency.⁴

The probability π may be used to compute the probability \hat{p} , that an entry (of 1 or 0) appearing in the reduced matrix reflects the subject's latent preference. Because the cumulative frequencies of 0 and 1 were mapped into a 0 and those of 2 and 3 into a 1 in his reduced matrix, the probability \hat{p} that the individual's "true" latent preference will appear in the reduced matrix is the probability that he chose according to his latent preference at least twice. It is, in other words, the probability that he made his latent choice on all three or that he made it only on two replications in any of the $\binom{3}{2} = 3$ ways. Hence, we may write:

$$\hat{p} = \pi^3 + 3\pi^2(1 - \pi).$$

The solution to this equation for each individual provides an estimate of his level of consistency based on replicating pairwise choices with no reference to any expectation theory.

Now if a person's latent preferences exactly satisfied a particular theory there would still be some inconsistency in his manifest choices. If the degree of consistency calculated under a particular theory, \hat{p} , is much smaller than that indexed by \hat{p} (which assumes that there is a constant probability of correspondence between choice and preference), it indicates that his behavior fits the theory rather poorly. If, on the other hand, his inconsistency is not increased by the effort to constrain his data to fit the

⁴ I.e., $\pi > 0.50 + 2\sigma$, where $\sigma^2 = pq/n$, $p = q = \frac{1}{2}$, $n = 32$.

theory, then so far as we can judge in the light of his manifest preference-inconsistency, his data fit the theory as well as they fit the theory that there is a true constant value of which \hat{p} is the estimate.

The manner⁵ in which \hat{p} may be estimated under a particular theory will be described in detail for EU theory; appropriate modifications for SEV and SEU theory will then be pointed out.

Under EU theory, as a consequence of Theorem 3, the only admissible latent patterns are the two compatible patterns $\begin{vmatrix} 0 \\ 0 \end{vmatrix}$ and $\begin{vmatrix} 1 \\ 1 \end{vmatrix}$ and the confirming pattern $\begin{vmatrix} 0 \\ 1 \end{vmatrix}$. In the manifest data, because of the stochastic properties of choice, the violating pattern $\begin{vmatrix} 1 \\ 0 \end{vmatrix}$ may be expected to occur as well. As \hat{p} gives the probability that a manifest choice will reflect a latent preference under EU theory, we may write down, as in Table 2, the probability of each pattern occurring in the manifest data conditional on which latent pattern obtains, (where $\tilde{q} = 1 - \hat{p}$).

TABLE 2

PROBABILITY OF THE COLUMN PATTERN BEING MANIFEST
IF THE ROW PATTERN IS LATENT

	$\begin{vmatrix} 0 \\ 0 \end{vmatrix}$	$\begin{vmatrix} 1 \\ 1 \end{vmatrix}$	$\begin{vmatrix} 0 \\ 1 \end{vmatrix}$	$\begin{vmatrix} 1 \\ 0 \end{vmatrix}$
$\begin{vmatrix} 0 \\ 0 \end{vmatrix}$	\hat{p}^2	\tilde{q}^2	$\hat{p}\tilde{q}$	$\tilde{p}\tilde{q}$
$\begin{vmatrix} 1 \\ 1 \end{vmatrix}$	\tilde{q}^2	\hat{p}^2	$\tilde{p}\tilde{q}$	$\hat{p}\tilde{q}$
$\begin{vmatrix} 0 \\ 1 \end{vmatrix}$	$\tilde{p}\tilde{q}$	$\tilde{p}\tilde{q}$	\hat{p}^2	\tilde{q}^2

Let θ be the proportion of latent compatible patterns for the individual and hence $1 - \theta$ the proportion of latent confirming patterns. The probability A of obtaining a compatible pattern, $\begin{vmatrix} 0 \\ 0 \end{vmatrix}$ or $\begin{vmatrix} 1 \\ 1 \end{vmatrix}$, in the individual's reduced matrix is

$$A = (\hat{p}^2 + \tilde{q}^2)\theta + 2\tilde{p}\tilde{q}(1 - \theta). \tag{10}$$

The probability B of a confirming pattern $\begin{vmatrix} 0 \\ 1 \end{vmatrix}$ being manifest is:

$$B = \tilde{p}\tilde{q}\theta + \hat{p}^2(1 - \theta). \tag{11}$$

And clearly, the probability of a violation $\begin{vmatrix} 1 \\ 0 \end{vmatrix}$ being manifest is $1 - A - B$.

⁵ We are indebted to Lincoln Moses for this suggestion.

If we let r , s , and t represent the number of compatible, confirming, and violating patterns manifest in a particular individual's reduced matrix, then the method of moments yields estimates of \hat{p} and θ by equating

$$A \text{ to } r/N \text{ and } B \text{ to } s/N,$$

where $N = r + s + t = 4\binom{8}{2} = 112$, that is the total number of EU patterns in the 8×4 reduced matrix showing $\binom{8}{2}$ patterns in each of its four columns. To be more specific, the method of moments estimates \hat{p} and θ by equating A to the proportion of compatible and B to that of confirming manifest patterns, substituting the obtained values for A and B in the Eqs. 10 and 11, and solving these for \hat{p} and θ simultaneously. These calculations are tedious and were done on an electronic computer. It may be remarked that exactly the same estimates would result from writing down the "likelihood"—as if all patterns were statistically independent, which they cannot be—and applying the method of maximum likelihood, in which case we would have

$$L(\hat{p}, \theta) = A^r B^s (1 - A - B)^t. \tag{12}$$

Identically, the same procedure can be carried out for SEV theory except, of course, that the values of r , s , and t are different, being based on row patterns rather than column patterns. Here, $r + s + t = \binom{2}{8} = 16$, as there are only two patterns in each of the eight rows.

The procedure for estimating \hat{p} and θ under SEU theory is essentially the same but slightly more complex. The patterns involved in testing SEU theory are the patterns in the second-order minors of the reduced data matrix. Every second-order minor indicates the order relation on the expectations of each of the four pairs of gambles, designated in the notation of Theorem 5 as $\begin{vmatrix} B & D \\ A & C \end{vmatrix}$. Theorem 5 says that if the entries in the minor were numerical ratios of expectations, rather than just order relations, the determinant of every second-order minor would equal zero, i.e., $AD = BC$.

The total number of possible patterns of second-order minors in the stochastic dominance matrix is $2^4 = 16$. If SEU theory holds, 14 of these 16 are admissible as true latent patterns, two are inadmissible or violating. The violating patterns are $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$ and $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$. Clearly, Theorem 5 is violated if A and D are both greater than 1 (or less than 1) and B and C are both less than 1 (or greater than 1). The remaining patterns may be collapsed into two classes, as in the case of EU and SEV theory, and again called compatible and confirming.

The motivation for using these terms for the classes is as follows: A pattern is classified as *compatible* if, given any three elements of the pattern, there is no constraint on the fourth element. The remaining patterns are classified as *confirming* or *violating* because, given one particular subset of three elements of the pattern, the fourth element is predicted by SEU theory and hence is either confirmed or violated. For example, consider the incomplete pattern $\begin{vmatrix} 1 & 0 \\ & \end{vmatrix}$, i.e., $B > 1, C > 1, D > 1$ (see Eq. 7). For the

determinant of this minor to equal zero, A must be greater than 1. Hence, the occurrence of a *one* is confirming, the occurrence of a *zero* is a violation.

The patterns are enumerated according to class in Table 3.

TABLE 3

CLASSIFICATION OF PATTERNS IN SECOND-ORDER MINORS ACCORDING TO SEU THEORY

Compatible	1 1	0 0	1 0	0 1	1 1	0 0			
	0 0	1 1	1 0	0 1	1 1	0 0			
Confirming	1 0	0 1	1 1	1 1	0 1	1 0	0 0	0 0	
	1 1	1 1	0 1	1 0	0 0	0 0	1 0	0 1	
Violating	0 1	1 0							
	1 0	0 1							

The equations for SEU theory that correspond to Eqs. 10 and 11 for EU and SEV theory are as follows:

$$A = (\hat{p}^4 + 4\hat{p}^2\hat{q}^2 + \hat{q}^4)\theta + 3(\hat{p}^3\hat{q} + \hat{p}\hat{q}^3)(1 - \theta), \tag{13}$$

$$B = 4(\hat{p}^3\hat{q} + \hat{p}\hat{q}^3)\theta + (\hat{p}^4 + 6\hat{p}^2\hat{q}^2 + \hat{q}^4)(1 - \theta). \tag{14}$$

These equations were obtained by writing out a 14×16 matrix like Table 2 for the manifest and latent SEU patterns and following the same procedure as was used to get Eqs. (10) and (11).

The expression for the likelihood of the data under SEU theory is exactly the same as Eq. (12) with the appropriate substitutions of A and B and where $r + s + t = \binom{8}{2} \binom{4}{2} = 168$, the total number of second-order minors.

It is important to note that for SEV theory the 16 tests contained in each matrix are experimentally independent whereas the 112 tests of EU theory and the 168 tests of SEU theory are not. The test of EU theory, however, is at least as powerful as the test of SEV theory, and the test of SEU theory is at least as powerful as that of EU theory.

For EV theory, the test is simpler. In the 8×4 reduced matrix there are 30 predictions of either a 1 or 0, as two of the cells represent judgments on pairs of gambles which have equal expected values (see Table 1). The expectation and variance of the number of times these predictions will fail, given \hat{p} , is

$$E(V) = 30(1 - \hat{p}), \tag{15}$$

$$\text{Var}(V) = 30\hat{p}(1 - \hat{p}). \tag{16}$$

Bidding Data

Only a single response (judgment) was obtained from each subject for each of the bidding conditions: buy, sell, and price; and so there was no control of inconsistency. Also, subjects tended to respond in multiples of 5 cents and 10 cents. Our intent had been to convert the bids of each kind into a (1, 0) data matrix of pair comparisons implied by the bids, realizing that an intermediate category of judgment was added—"indifference". Adding this intermediate category of response, of course, multiplies the variety of response patterns under each model. This fact, in conjunction with the lack of control on inconsistency, discouraged us from devising a test of the theories using the bidding data. Instead, the data were pooled for various subgroups and analyzed for trends which might be useful for guidance in future experiments using this method of data collection.

C. RESULTS

There was one subject who did not complete his tasks during the experimental session (Subject 111, WVM). As his data were not analyzed, the following discussion deals with the 39 remaining subjects.

Inconsistency of Judgment

A value of π was obtained for each subject as a solution to Eq. 9. There were three subjects who failed to show significant levels of consistency. It is an interesting observation that all three of these subjects had values of $\lambda < \frac{1}{4}$ —they had fewer 0's and 3's than are to be expected by tossing a fair coin—they were more inconsistent than pure chance. This could be an unusual statistical event with the subjects' having no true preference or, what is perhaps more reasonable, these subjects changed their basis of preference in the course of the experiment. In any event, no parameter estimates could be made for these three subjects.

The distribution of π values for the remaining subjects is given in Table 4.

An arcsine transformation was applied on these π values and a two-way analysis of variance made to compare College Students with Willow Village subjects and Males and Females. The results are reported in Table 5. The significant interaction reflects the fact that college student males tended to be highly consistent and Willow Village males much less so, whereas, there is little difference between the two groups of women.

Tests of Expectation Models

The specific results for each subject under each theory are presented in Table 6.

The four theories form a partial order, with SEU the most general case, EU and SEV special cases but distinct, and EV theory a special case of each of them. If a

TABLE 4
DISTRIBUTION OF π VALUES, EXPERIMENT I.

π	WVM	WVF	CSM	CSF
0.70	1			
0.71				
0.72				
0.73				1
0.74				
0.75	1			
0.76				
0.77	1	1		
0.78				
0.79	1			1
0.80				
0.81				1
0.82				
0.83				
0.84				1
0.85	1	1	1	
0.86				
0.87				1
0.88				
0.89				
0.90		1		2
0.91				2
0.92			1	1
0.93				1
0.94				
0.95		1	1	
0.96		1		1
0.97	1	1	1	
0.98	1			1
0.99			3	
1.00		1	2	
Total	7	7	9	13

TABLE 5
ANALYSIS OF INCONSISTENCY

Source	<i>df</i>	<i>MS</i>	<i>F</i>
CS vs WV (A)	1	192.09	2.984
SEX (B)	1	14.69	—
A \times B	1	726.99	11.294 ^a
Error	32	64.37	

^a Significant at 0.01 level.

TABLE 6
PARAMETER ESTIMATES UNDER EACH THEORY FOR EACH SUBJECT, EXPERIMENT I^a

Subj.	EV				EU		SEV		SEU	
	\hat{p}_i	Exp.# Viol.	Obs.# Viol.	Signif. Beyond	\hat{p}_i	θ_i	\hat{p}_i	θ_i	\hat{p}_i	θ_i
101	—	—	—	—	—	—	—	—	—	—
102	0.885	3.4	8	0.010	0.79	0.66	0.87	0.50	0.94	0.46
104	0.999	0.1	15	0.001	—	—	—	—	—	—
105	0.942	1.7	7	0.001	0.95	0.56	0.90	0.69	0.94	0.58
110	0.866	4.0	6	—	0.96	0.44	0.99	0.88	0.999	0.76
113	0.789	6.3	7	—	0.84	0.41	0.84	0.91	0.86	0.49
114	—	—	—	—	—	—	—	—	—	—
115	0.844	4.7	11	0.001	0.78	0.55	0.94	0.99	0.74	0.96
117	0.997	0.1	15	0.001	—	—	—	—	—	—
103	1.000	0.0	15	0.001	—	—	—	—	—	—
106	0.997	0.1	15	0.001	0.999	0.98	0.99	0.95	0.999	0.88
107	0.991	0.3	8	0.001	0.99	0.51	—	—	0.999	0.83
108	0.866	4.0	11	0.001	0.90	0.91	0.92	0.85	0.90	0.74
109	0.994	0.2	8	0.001	0.99	0.42	—	—	—	—
112	0.942	1.7	13	0.001	0.99	0.88	—	—	0.999	0.83
116	0.969	0.9	8	0.001	0.99	0.64	0.99	0.82	0.999	0.73
118	—	—	—	0.001	—	—	—	—	—	—
203	0.991	0.3	8	0.001	0.99	0.51	0.94	0.99	0.999	0.83
204	1.000	0.0	0	—	0.99	0.61	0.99	0.43	0.999	0.59
206	1.000	0.0	8	0.001	0.99	0.57	—	—	—	—
207	0.997	0.1	11	0.001	0.99	0.75	—	—	—	—
210	1.000	0.0	15	0.001	—	—	—	—	—	—
213	0.982	0.5	15	0.001	—	—	—	—	—	—
219	1.000	0.0	8	0.001	0.99	0.46	—	—	—	—
220	1.000	0.0	15	0.001	—	—	—	—	—	—
222	0.942	1.7	15	0.001	0.99	0.95	0.999	0.19	0.999	0.88
201	0.976	0.7	8	0.001	0.99	0.46	—	—	—	—
202	0.969	0.9	4	0.001	0.99	0.48	0.99	0.75	0.999	0.67
205	0.819	5.4	12	0.001	0.97	0.73	0.93	0.93	0.90	0.93
208	0.982	0.5	7	0.001	0.96	0.45	0.99	0.95	0.94	0.89
209	0.930	2.1	13	0.001	0.62	0.96	0.99	0.82	0.94	0.64
211	0.952	1.4	8	0.001	0.98	0.57	0.99	0.82	0.999	0.73
212	0.976	0.7	11	0.001	0.99	0.75	—	—	—	—
214	0.969	0.9	7	0.001	0.95	0.51	0.84	0.91	0.98	0.55
215	0.902	2.9	8	0.001	0.96	0.47	0.93	0.93	0.96	0.76
216	0.994	0.2	12	0.001	0.99	0.82	0.99	0.95	0.999	0.88
217	0.999	0.1	15	0.001	—	—	—	—	—	—
218	0.987	0.4	8	0.001	0.96	0.44	0.94	0.99	0.999	0.76
221	0.885	3.4	9	0.001	0.99	0.66	0.99	0.88	0.999	0.83

^a A dash "—" indicates no parameter estimates could be made in that instance, e.g. \hat{p}_i is imaginary, or the data matrix consists of all 1's or 0's and hence no information in it, etc.

particular theory is found not to hold for a subject, then it follows that any special cases of it would not hold either. To reject SEU theory for some subject, for example, is to reject all the other theories for that subject.⁶

SEU theory could be tested on 23 of the subjects and rejected for subjects 115 and 208.

SEV theory could be tested on 21 of the subjects and rejected for the following seven subjects: 102, 105, 115, 203, 208, 214, 218. This theory is rejected for 115 and 208 because SEU theory is rejected for them.

EU theory could be tested on 29 of the subjects and rejected for the following seven subjects: 102, 115, 206, 208, 209, 214, 218. This theory is rejected for 115 and 208 because SEU theory is rejected for them.

It comes as no surprise that EV theory may be rejected for 34 of 37 subjects on whom it could be tested. Only for Subjects 110, 113, and 204 could it not be rejected.

Bidding Data

For each of the subgroups listed in Table 7, the average bid for each of the 64 gambles was computed under each of the bidding instructions. Slope (m), intercept (k) and correlation (r) between average bid and EV are reported in the table.

TABLE 7
SLOPE (m), INTERCEPT (k), AND CORRELATION (r), BETWEEN AVERAGE BID
AND EXPECTED VALUE

Group	buying bids			selling bids			price bids		
	m	k	r	m	k	r	m	k	r
Willow Village Men	0.22	0.04	0.76	0.51	0.12	0.96	0.49	0.09	0.98
Willow Village Women	0.20	0.07	0.94	0.60	0.39	0.88	0.49	0.40	0.88
All Willow Village	0.22	0.05	0.88	0.56	0.25	0.95	0.49	0.23	0.96
College Student Men	0.72	0.01	0.98	0.88	0.88	0.99	0.87	0.04	1.00
College Student Women	0.31	0.04	0.96	0.77	0.77	0.99	0.68	0.05	0.99
All College Students	0.48	0.02	0.98	0.81	0.81	0.99	0.76	0.05	0.99
All Subjects	0.36	0.04	0.97	0.70	0.70	0.99	0.64	0.13	0.99

⁶ Differences in the third decimal place (e.g., Subject 105 for SEU theory) were not counted against a theory nor were differences due to rounding off error (e.g., Subjects 204 and 219 for whom $\hat{p}_1 < 1.000$ if the fourth decimal place had been reported).

Figures 3 and 4 portray the relations for the buying and selling data separately for the Willow Village subjects and the College students. The plots for the price bids yields regression lines intermediate between those for buy and sell bids but rather closer to that for selling bids. This holds for each group of subjects. In general, our subjects bid more than the expected value for gambles with low expectations and underbid gambles whose expected values were higher, but the overall tendency was clearly to underbid.

4. EXPERIMENT II

Experiment I was essentially a preliminary trial of the general experimental design and the method of analysis. Experiment II is a much more substantial study in terms of the number of stimuli, the number of subjects, the realism of the experimental situation, and the involvement of the subjects. The collection and analysis of these data were under the supervision of the third author.

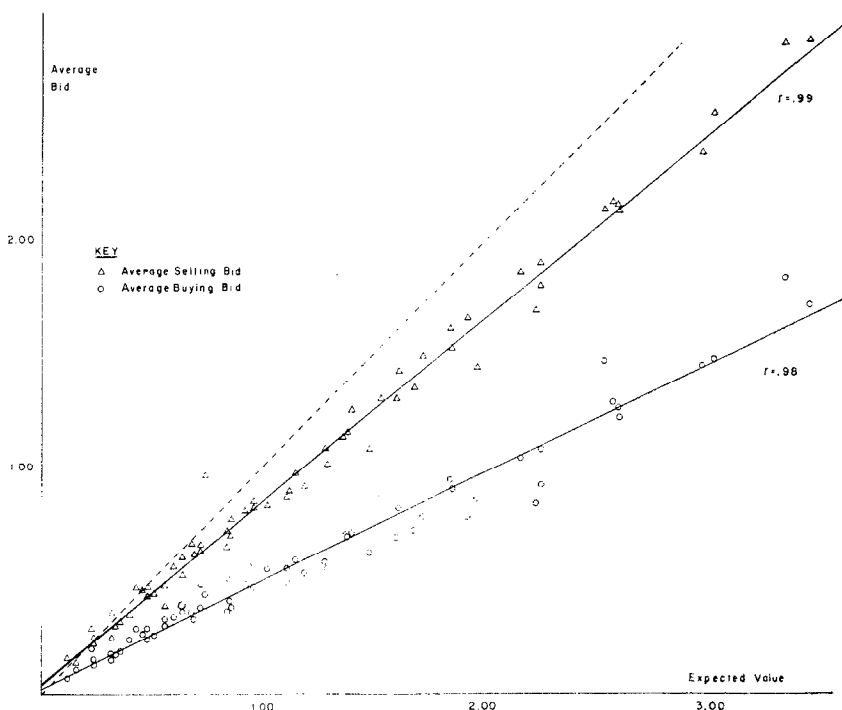


FIG. 3. Relation of bids to expected value for all Willow Village subjects.

A. EXPERIMENTAL DESIGN

Stimuli

Figure 5 displays the 47 distinct gambles used in Experiment II, and their organization into six sets with six pairs of gambles in each set. This design permits the probabilities to have about the same range as in Experiment I and yields a 6×6 data matrix; hence the same number of tests can be made of SEV theory as of EU theory, whereas in Experiment I the data matrix was 8×4 . The gambles are of the form used in Experiment I and, of course, the theorems pertain to them: some finite probability of a finite prize, otherwise nothing. In this experiment real gambles were used, and each prize was some number of cigarettes.⁷

The expected value of each gamble is contained in Table 8. The configuration of cells in Table 8 corresponds to the configuration of gambles as portrayed in Fig. 5. So, for example, the top pair of gambles in Set I, [(7/8, 5), (6/8, 10)] as shown in Fig. 5 may be seen to have expected values [(4.38), (7.50)].

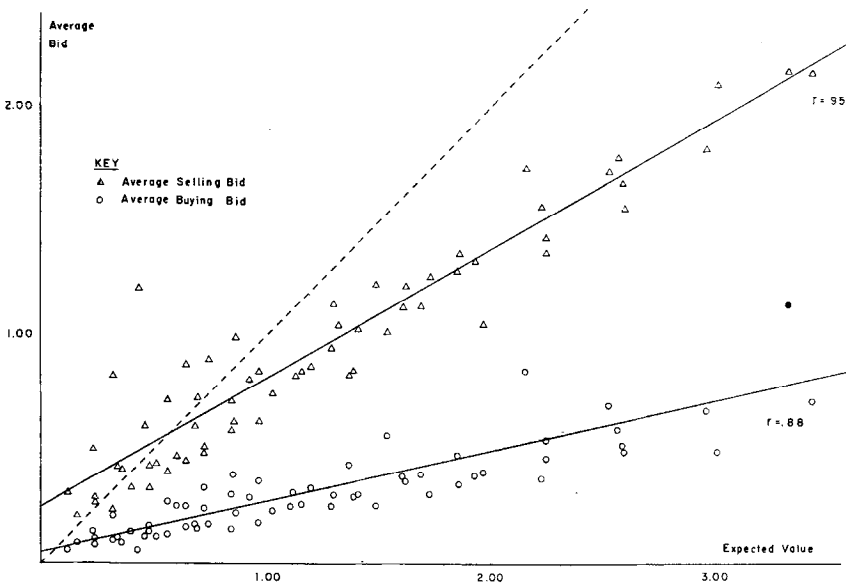


FIG. 4. Relation of bids to expected value for all college students.

⁷ Pall Mall were the particular cigarettes used. Glassine stamp envelopes were used to package lots of five cigarettes. These were used as payoffs, alone, or as supplements to a normal package of 20 as required.

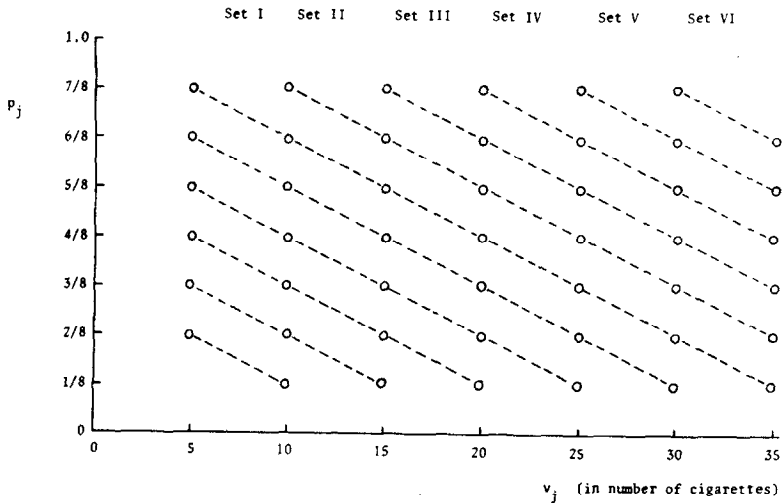


FIG. 5. The six sets of pairs of gambles in Experiment II.

TABLE 8
EXPECTED VALUES OF GAMBLES USED IN EXPERIMENT II

Prob. ^a	Number of Cigarettes to be Won						
	5	10	15	20	25	30	35
7/8	4.38	8.75	13.12	17.50	21.88	26.25	—
6/8	3.75	7.50	11.25	15.00	18.75	22.50	26.25
5/8	3.12	6.25	9.38	12.50	15.62	18.75	21.88
4/8	2.50	5.00	7.50	11.00	12.50	15.00	17.50
3/8	1.88	3.75	5.63	7.50	9.38	11.25	13.12
2/8	1.25	2.50	3.75	6.00	6.25	7.50	8.75
1/8	—	1.25	1.88	2.50	3.12	3.75	4.38

^a Indicates probability of winning.

Each gamble was set up on a wheel of fortune and photographed with the spinner removed. The wheel of fortune had eight circles around the rim (instead of 10 as portrayed in the booklets used in Experiment I). From 1 to 7 of the circles was colored black to indicate the probability of positive outcome, and these black circles were oriented symmetrically about the top of the wheel of fortune. These photographs were then used to construct slides of the pair comparisons. Hence these stimuli will be called the "slide" stimuli to distinguish them from the "booklet stimuli" used in Experiment I.

Subjects

The subjects were inmates of the State Prison of Southern Michigan, Jackson, Michigan, a very large, maximum security prison. Initially, 120 men were selected from volunteers responding to an article in the prison newspaper. Those who were due for parole or had IQ's below 90 were rejected, as well as those for some other sufficient and prohibitive cause.

Procedure

The 120 subjects were divided into six convenient groups of 20 each for an experimental session. An experimental session was scheduled for 2 hours. Each group of men undertook five experimental sessions at 1-week intervals followed by additional sessions in which related data were collected—some of which are reported here. The experimenters visited the prison on the same two successive days each week, running three groups of men each day.

A group of subjects went through a single replication of the 36 pair comparisons in each experimental session. Each pair of gambles was projected on a screen and each subject marked his preference on a data sheet. The 36 pairs were randomized in two different orders, each used in alternate sessions one week apart. The gamble technically defined as the *left* gamble in Fig. 1 *appeared* on the left in one order and on the right in the other order.

The experimental slides were alternated with color slides of scenery, etc., which were left on for $\frac{1}{2}$ to 3 minutes, depending on how much the men liked it and wanted to discuss it. At the beginning of the first experimental session the experimental slides were exposed for 1 minute per slide and by the fifth session the exposure duration had been reduced to 15 seconds. In the latter sessions, the time available (about $\frac{3}{4}$ hours) was used to collect additional data.

The subjects were informed that on each day at the end of the 36 experimental slides, one would be chosen to be played for real, and that each subject would play the particular gamble he had preferred in that pair. Then at the end of the 36 pairs, the data sheets were collected, the subjects filed up, and the pair selected was projected on the screen. Two wheels of fortune were set up on a table beneath the screen corresponding to the two gambles portrayed; and as each subject filed up to the table, his choice was called out by an assistant who read it from the subject's data sheet. The subject then spun the wheel of fortune that he had chosen and was immediately paid off in cigarettes if he won, otherwise nothing.

The pair of gambles chosen to be played was selected to be *nearly* equal in EV so that there would be less likelihood of a consensus of choice influencing an individual's later choices. A different pair was chosen for each group of subjects run on the same day.

Over the five sessions a subject's expected total winning was approximately 60

cigarettes. In addition each subject was paid \$0.50 per experimental session if he completed all five sessions, otherwise \$0.25 per session for those he completed. Out of the initial 120 subjects, 108 completed the first five sessions.

Further experimental sessions were run, again at 1-week intervals, in which additional data were collected on those subjects who volunteered to continue. Some of the additional data collected were on Booklets 1, 3, and 5 used in Expt. I. Booklet 1 was administered in Session 6 and Booklets 3 and 5 were administered at the beginning and the end of a later session with intervening tasks lasting about an hour. The procedure followed was otherwise the same. Seventy-eight subjects finished these booklets.

B. METHOD AND ANALYSIS

The method of analysis for the pair comparisons data on the new stimuli was essentially the same as that for Booklets 1, 3, and 5 in Experiment I, except for some alterations to adapt to five replications and to six sets of stimuli.

In Experiment I analysis, we estimated π on the basis of the number of perfectly consistent responses (0's and 3's) in the cumulative total matrix. A similar approach could have been used in Experiment II (computing π from the number of 0's and 5's in the cumulative total matrix over five replications); but it would have required the solution of 5th-degree equations, so another approach was used which appeared to be simpler for our purposes. The computer formed all subsets of three replications and computed a λ and π on each, using the same formulas as for Experiment I (Eqs. 8 and 9). Each π computed from a subset of three replications is an unbiased estimate of the individual's latent consistency. Therefore if the π 's from the subsets are averaged, the resulting average π will likewise be an unbiased estimate. This average π has the same interpretation mentioned previously, namely the probability that a subject's manifest choice on any particular occasion is identical to his latent preference, without assuming anything about any expectation theories.

This π may be used as before to compute an index of consistency \hat{p} for the individual's reduced matrix. The index \hat{p} is the probability that a stochastically dominant choice of 0 or 1 (reduced from five replications) is the same as the subject's latent preference. They will be the same only if he makes his latent choice on more than half of the five replications. The formula for \hat{p} is therefore based on the number of ways that this can happen:

$$\hat{p} = \pi^5 + 5\pi^4(1 - \pi) + 10\pi^3(1 - \pi)^2.$$

There is no essential change in the method for estimating the \hat{p} indices of consistency based upon the EU, SEV, and SEU expectation theories, since neither the dimensions of the matrix nor the number of replications entered into the derivation of the \hat{p} formulae (Eqs. 10-14). The only change was in the number of predictions for

each theory as follows: $N = 90$ for EU and SEV theory, and $N = 225$ for SEU theory.

The method of analysis for EV theory was unchanged from Experiment I. For a subject who maximizes expected value, predictions can be made of his paired comparison choices by referring to Table 8. The two gambles from any pair formed along the diagonal from the lower left to the upper right have equal expected value, so no predictions can be made for those six pairs. EV predictions can be made for the remaining 30 pairs, however, with 0 predicted for all pairs above the diagonal and 1 for all pairs below the diagonal.

C. RESULTS

We will summarize the results with primary attention to the results on the slide stimuli, since this part of Experiment II was highly realistic and the subjects were much involved.

Inconsistency of judgment

The distribution of π values is given in Table 9. There were two subjects (058 and 082) out of the 108 whose level of inconsistency was within 2σ of chance. Interestingly enough, when they later responded to the booklet stimuli they were quite consistent. Conversely, two subjects who were consistent on the slide stimuli (014 and 033) were the most inconsistent on the booklet stimuli.

No statistical comparison of these distributions has been made with the college student or Willow Village population because of the difference in the stimuli in the case of slides against booklets and because of the difference in procedure in the case of the booklet stimuli. In the latter case the second and third replications were a week after the first in Experiment II and the subjects were highly experienced after five previous sessions on the slide stimuli.

Test of Expectation Models

We present in Table 10 a summary of the results on testing the expectation models with the slide stimuli of Experiment II. Of the 87 subjects for whom EV theory was rejected, 72 deviated at the 0.001 level, and seven at the 0.01 level. EV theory was rejected for eight other subjects because they violated one or more of the other theories of which EV theory is a special case.

For the other three theories, the results were as follows. SEU theory could be tested on 85 subjects and rejected⁸ for four of them (5%).

⁸ A theory was rejected only if $\hat{p} - \bar{p} \geq 0.010$, to be consistent in dealing with rounding the effects.

TABLE 9
DISTRIBUTION OF π VALUES, EXPERIMENT II

π	Booklet Stimuli	Slide Stimuli
0.40	1	-
0.56	1	-
0.58	-	1
0.66	-	1
0.68	2	-
0.70	1	3
0.71	-	2
0.72	-	-
0.73	2	1
0.74	-	1
0.75	1	1
0.76	-	2
0.77	1	3
0.78	-	1
0.79	2	-
0.80	-	4
0.81	2	3
0.82	-	4
0.83	-	1
0.84	4	-
0.85	1	6
0.86	-	5
0.87	4	-
0.88	6	6
0.89	-	6
0.90	4	9
0.91	9	5
0.92	4	10
0.93	2	4
0.94	8	5
0.95	-	6
0.96	6	4
0.97	3	4
0.98	4	2
0.99	2	4
1.00	8	4

TABLE 10
RESULTS ON TESTING THE EXPECTATION MODELS WITH THE
SLIDE STIMULI OF EXPERIMENT II

No. of Subjects	EV	EU	SEV	SEU
Theory Tested	108	92	85	85
Theory Rejected	87	23	29	4

SEV theory could be tested on 85 subjects and rejected for 29 of them (34%). This theory is rejected for two of the 29 because SEU theory is rejected for them.

EU theory could be tested on 92 subjects and rejected for 23 of them (25%). This theory is rejected for one of the 23 because SEU theory is rejected for him.

The results on testing the expectation models with the booklet stimuli are summarized in Table 11 which includes the results from Experiment I for comparison. It is

TABLE 11
RESULTS ON TESTING THE EXPECTATION
MODELS WITH THE BOOKLET STIMULI

Expt. No.	No. of Subjects	EV	EU	SEV	SEU
II	Theory Tested	78	66	43	52
	Theory Rejected	67	15	8	5
I	Theory Tested	37	29	21	23
	Theory Rejected	34	7	7	2

evident that the results on the booklet stimuli are quite similar in the two experiments and they differ little from the results on the slide stimuli, and mostly in testing SEU theory. For the booklet stimuli and under the hypothetical gambling conditions used, SEU theory could be rejected for about 10% of the subjects in both Experiments I and II and with more realism and high involvement, SEU theory was still rejected for 5% of the subjects.

ACKNOWLEDGMENT

We have had the advantage of discussing some of the statistical aspects of this paper with L. J. Savage, John Tukey, Lincoln Moses, and Robyn M. Dawes, but there is no implication that they concur in the procedures we used. We also wish to thank the Department of Corrections of the State of Michigan, particularly Mr. Gus Harrison, Director, and Dr. Fred Pesetsky, Chief Psychologist of the Psychiatric Clinic, and also the administration of the State Prison of Southern Michigan, particularly Mr. George A. Kropp, Warden, for their cooperation and assistance in collecting a substantial part of the experimental data analyzed here.

REFERENCES

- BERNOULLI, D. Exposition of a new theory on the measurement of risk. *Econometrica*, 1954, 22, 23-36.
- CHERNOFF, H. AND MOSES L. *Elementary decision theory*. New York: Wiley, 1959.
- COOMBS, C. H. AND KOMORITA, S. S. Measuring utility of money through decisions. *American Journal of Psychology*, 1958, 71, 383-389.
- DAVIDSON, D. AND MARSCHAK, J. Experimental tests of a stochastic decision theory. In C. W. Churchman and P. Ratoosh (Eds.) *Measurement: Definitions and theories*. New York: Wiley, 1959.
- DAVIDSON, D., SUPPES, P., AND SIEGEL, S. *Decision making: An experimental approach*. Stanford, California: Stanford University Press, 1957.
- EDWARDS, W. The theory of decision making. *Psychological Bulletin*, 1954, 51, 380-417.
- EDWARDS, W. The prediction of decisions among bets. *Journal of experimental Psychology*, 1955, 50, 201-214.
- EDWARDS, W. Behavioral decision theory. *Annual Review of Psychology*, 1961, 12, 473-498.
- HURST, P. M. AND SIEGEL, S. Prediction of decisions from a higher ordered metric scale of utility. *Journal of experimental Psychology*, 1956, 52, 138-144.
- MOSTELLER, R. AND NOGEE, P. An experimental measurement of utility. *Journal of Political Economics*, 1951, 59, 371-404.

RECEIVED: March 18, 1966.