ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 71, No. 2, August, pp. 161–194, 1997
ARTICLE NO. OB972721

Tests of Theories of Decision Making: Violations of Branch Independence and Distribution Independence

Michael H. Birnbaum

California State University, Fullerton, and Institute for Mathematical Behavioral Sciences, Irvine

and

Alfredo Chavez

California State University, Fullerton

This paper investigates two behavioral properties of choices between gambles: branch independence and distribution independence. Branch independence is the assumption that probability-outcome branches that are identical in two gambles being compared should have no effect on preference order. In the test of distribution independence, two middle outcomes of fixed probability are sandwiched between two extreme outcomes, common to both gambles, whose probabilities are varied. Distribution independence requires that preference order should be independent of the probability distribution of these common outcomes. Distribution independence is implied by original prospect theory and the configural weight model of Birnbaum and McIntosh (1996). For both properties, patterns of violations are opposite those predicted by the inverse-S weighting function used in the model of cumulative prospect theory by Tversky and Kahneman (1992).

Key Words: Allais paradox; branch independence; comonotonic independence; cumulative prospect theory; configural weighting; distribution independence;

We thank Robert Bailey and Heather Orey for assistance. Thanks are also due to R. Duncan Luce for his suggestions on an earlier draft, and to two anonymous reviewers. Support was received from National Science Foundation Grant SBR-9410572 to the first author through California State University, Fullerton.

Address reprint requests to Michael H. Birnbaum, Department of Psychology, California State University, P. O. Box 6846, Fullerton, CA 92834-6846. E-mail address: mbirnbaum@fullerton.edu.

0749-5978/97 $25.00

Ellsberg paradox; expected utility theory; independence; interval independence; rank dependent utility; Savage's axiom; subjective expected utility; sure-thing principle; outcome independence; weighting function.

In recent years, investigators studying decision making have reached a consensus that Subjective Expected Utility (SEU) theory (Savage, 1954) is not descriptive of empirical choices, and that a generalization of the SEU model is required to account for human behavior (Birnbaum, 1992; Edwards, 1992; Luce, 1992; Lopes, 1990; Stevenson, Busemeyer, & Naylor, 1991; Tversky & Kahneman, 1992). Several models have been proposed in which the weight of an outcome depends (at least in part) on the rank of the outcome among the other possible outcomes of a gamble, as well as the outcome's probability. Although these new theories can make the same predictions in certain experiments, they are distinct and make different predictions in experiments crafted to test between them.

This paper examines two properties of decision behavior and analyzes them with respect to different models of decision making. The first property is a weaker form of Savage's "sure thing" axiom called *branch independence*. The second property, *distribution independence*, is implied by the configural weight model of Birnbaum and McIntosh (1996) and original prospect theory (Kahneman & Tversky, 1979), but can be violated in cumulative prospect theory (Tversky & Kahneman, 1992; Luce & Fishburn, 1991, 1995) and Birnbaum and Stegner's (1979) configural weight model.

Recent experiments with judgments and choices among three-outcome gambles found evidence of systematic violations of branch independence. Let $(x, p; y, q; z, r)$ represent a three-outcome gamble, in which the probabilities of receiving x, y, or z are p, q, and $r = 1 - p - q$, respectively. Branch independence requires that $(x, p; y, q; z, r)$ is preferred to $(x', p'; y', q'; z, r)$ if and only if $(x, p, y, q; z', r)$ is preferred to $(x', p', y', q'; z', r)$. Figure 1 illustrates the property

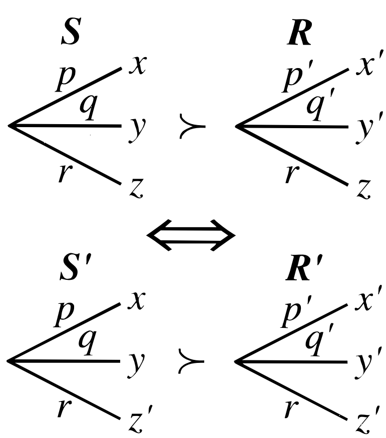

FIG. 1. Branch independence in three-outcome gambles: S is preferred to R if and only if S' is preferred to R'. Note that removing the common branch (z, r) or (z', r) leaves the same contrast in both cases.

of branch independence, $S \succ R$ if and only if $S' \succ R'$, where $\succ$ is the preference relation. Figure 1 illustrates that the outcome (z) on the common branch (z, r) can be changed to (z', r) without changing the preference order.

Birnbaum and McIntosh (1996) found systematic violations of branch independence in choices between gambles with three equally likely outcomes, denoted (x, y, z). Birnbaum and McIntosh (1996) found that the majority of people chose ($2, $40, $44) over ($2, $10, $98); however, the majority chose ($10, $98, $108) over ($40, $44, $108). This pattern of violation of branch independence was repeated with other sets of outcomes, it was representative of the results with different outcome values, and it was descriptive of individual judges' data as well as group proportions.

Because branch independence is weaker than Savage's "sure thing" axiom, violations observed by Birnbaum and McIntosh (1996) also violate Savage's (1954) axiom. Violations indicate that choices are not consistent with Expected Utility theory (EU), or Subjective Expected Utility (SEU) theory (Savage, 1954). They also are inconsistent with Subjectively Weighted Utility (SWU) theories (Edwards, 1954; Karmarkar, 1978; Viscusi, 1989). For example, suppose SWU $= \Sigma\ S(p_i)u(x_i)$. SWU implies branch independence for three outcome gambles, as in Figure 1, when all of the outcomes are distinct. Proof: SWU(S) $>$ SWU(R) if and only if $S(p)u(x) + S(q)u(y) + S(r)u(z) > S(p')u(x') + S(q')u(y') + S(r)u(z)$; subtracting $S(r)u(z)$ from both sides and adding $S(r)u(z')$ to both sides, we have, $S(p)u(x) + S(q)u(y) + S(r)u(z') > S(p')u(x') + S(q')u(y') + S(r)u(z')$; this relation holds if and only if SWU(S') $>$ SWU(R').

Systematic violations of branch independence in Birnbaum and McIntosh (1996) are also inconsistent with the theory that judges consistently cancel common components when choosing between gambles (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). If common branches were "trimmed," before comparison, there should be no systematic violations of independence. Note that in Fig. 1, removing the common branches leaves the same comparison in both cases.

The pattern of violations observed in these studies is opposite that implied by the inverse-S decumulative weighting function used by Tversky and Kahneman (1992) to fit certainty equivalents of binary gambles. Wu and Gonzalez (1996) concluded that cumulative prospect theory implies an inverse-S decumulative weighting function to explain choices between two- and three-outcome gambles. Whereas Tversky and Kahneman (1992) had used parametric assumptions to fit their cumulative weighting function, the technique of Wu and Gonzalez (1996) estimates the curvature of the weighting function without such assumptions.

Similar violations of branch independence were obtained with judgments of buying prices and selling prices of gambles composed of two, three, and four equally likely outcomes (Birnbaum & Beeghley, 1997; Birnbaum & Veira, in press). These results were fit by the theory that the utility function is invariant of task (judgment versus choice), and that configural weights depend on the task and the judge's point of view (Birnbaum, 1997; Birnbaum, Coffey, Mellers, & Weiss, 1992; Birnbaum & Sutton, 1992). *Point of view* refers to instructions, such as those to identify with the buyer or seller, that asymmetrically affect the costs of judgment or decision errors (Birnbaum and Stegner, 1979). The fact that the same pattern of violations of branch independence is observed in both judgment and choice suggests that violations are not due to a process, such as editing, that would be specific to choice and not judgment.

The experiments of Birnbaum and McIntosh (1996), Birnbaum and Beeghley (1997) and Birnbaum and Veira (in press), which showed substantial violations of branch independence, all used gambles in which the outcomes are equally likely, and probability was not explicitly stated on each trial. Experiments by Wakker *et al.* (1994) and Weber and Kirsner (1997), which found smaller violations, used gambles in which the probabilities were made explicit.

It is possible that with explicitly stated probabilities, judges might conform to branch independence or to violate it in the manner consistent with the inverse-S cumulative weighting function. Therefore, this study replicates and extends previous research by investigating violations of branch independence with three- and four-outcome gambles in which probabilities are explicitly displayed.

This study also investigates a previously untested implication of the weighting function used in cumulative prospect theory for a paradigm in which judges choose between pairs of gambles in which two intermediate outcomes have fixed probability, but the probabilities of extreme outcomes are changed, in order to vary the decumulative probabilities of the outcomes. This *distribution independence paradigm*, described in a later section, allows us to test whether the configural weight of an outcome depends purely on the rank of the value of that outcome and that outcome's probability, or if it depends on the cumulative probability of the outcome, as postulated by cumulative prospect theory.

In the next section, we analyze violations of branch independence in four-outcome gambles; in a following section, we analyze our new distribution independence paradigm. Two main theories are contrasted, Cumulative Prospect Theory (with various special cases of weighting functions), and the configural weight models of Birnbaum and McIntosh (1996) and Birnbaum and Stegner (1979). An experiment is presented to test branch independence and distribution independence.

THEORETICAL ANALYSIS

It is useful to analyze the present experiment with respect to a generic, rank-dependent configural weight model, of which the models to be compared are (in this experiment) special cases. This generic model allows the weight of an outcome to depend on the outcome's probability, the outcome's rank, and the distribution of probabilities in the gamble. It will be assumed that when the number of distinct outcomes and the probability distributions are the same in two gambles being compared, that common branches with common ranks (comonotonic, probability-event-outcome combinations) will have the same value in both gambles and can therefore be subtracted off of both sides of a comparison and replaced with another common branch of the same rank and probability.

Let $(x, p; y, q; z, r; v, 1 - p - q - r)$ represent a four-outcome gamble with nonzero probabilities of p, q, r, and $1 - p - q - r$ to win x, y, z, or v, respectively. In this experiment, the probability mechanism is described to the judges as follows: There is an urn containing 100 slips of paper, which are otherwise identical, except $100p$ of the slips have the value of x printed on them; $100q$ have y; $100r$ have z; and the rest have v printed on them. The slips will be thoroughly mixed, and the prize will be determined by the amount written on the slip drawn randomly from the urn.

Branch Independence

Let $\succ$ represent a preference relation between gambles. The property of branch independence that will be tested in this experiment is defined as follows:

$$(x, p; y, q; z, r; v, 1 - p - q - r) \succ (x', p'; y', q'; z, r; v, 1 - p' - q' - r)$$
$$\text{if and only if} \tag{1}$$
$$(x, p; y, q; z', r; v', 1 - p - q - r) \succ (x', p'; y', q'; z', r; v', 1 - p' - q' - r),$$

where $p + q = p' + q'$. This property assumes that the trade-off between $(x, p; y, q)$ and $(x', p'; y', q')$ should be independent of the outcomes of branches common to both gambles being compared, which have the same probabilities in all cases.

In this study, we test branch independence with four-outcome gambles with four equal probabilities ($p = p' = q = q' = r = 1 - p - q - r = 1/4$), and we also test it with three-outcome gambles, combining the outcomes z and v($p = p' = q = q' = 1/4$, $r = 1/2$). Figure 2 illustrates a test of branch independence with four-outcome gambles. Note that the two lowest outcomes are changed from lowest (z, v) to highest (z', v'), and if these branches were removed, the contrast between the gambles would be the same $[(x, y)$ versus $(x', y')]$.

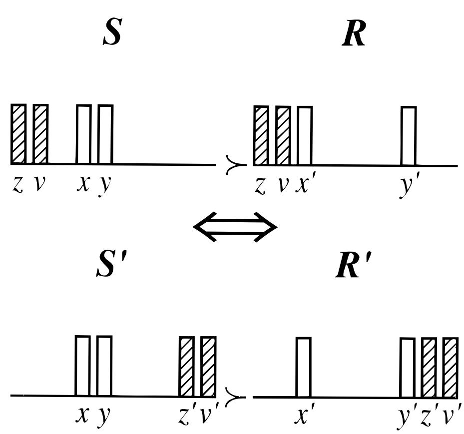

FIG. 2. A test of branch independence with four equally likely outcomes. Bars depict probabilities of outcomes; common outcomes are shown as shaded bars.

Generic Rank-Dependent Utility

The Generic Rank-Dependent Utility ($U(G)$) for such four-outcome gambles can be written as follows:

$$U(x, .25; y, .25; z', .25; v', .25) = w_L u(x) + w_{ML} u(y) + w_{MH} u(z') + w_H u(v') \qquad (2)$$

where $0 < x < y < z' < v'$; $u(x)$ is the utility function; w_L, w_{ML}, w_{MH}, and w_H are the relative, configural weights of the Lowest, Medium-Low, Medium-High, and Highest outcomes (at equal probabilities of $\frac{1}{4}$), respectively, which are assumed to be non-negative. We term Eq. (2) *generic* because it imposes no other restrictions on the weights, and thus does not distinguish configural weight models from cumulative prospect models.

Variations of this generic model have been analyzed in a number of papers (Birnbaum, 1974, 1982, 1992, 1997; Birnbaum & Jou, 1990; Birnbaum & Stegner, 1979, 1981; Birnbaum & Sotoodeh, 1991; Birnbaum & Zimmermann, submitted; Champagne & Stevenson, 1994; Lopes, 1990; Luce, 1992, 1996; Luce & Fishburn, 1991, 1995; Miyamoto, 1989; Quiggin, 1982; Starmer & Sugden, 1989; Tversky & Kahneman, 1992; Wakker, Erev, & Weber, 1994; Wakker & Tversky, 1993; Weber, 1994; Wu & Gonzalez, 1996). This experiment will compare some of these models that make different predictions for tests of branch independence and distribution independence.

In a choice between two of such gambles, assume that the judge chooses the gamble with the higher $U(G)$. For example, consider the following preference relation,

$$S' = (x, .25; y, .25; z', .25; v', .25) \succ R' = (x', .25; y', .25; z', .25; v', .25).$$

This relation ($S' \succ R'$) holds if and only if

$$U(x, .25; y, .25; z', .25; v', .25) > U(x', .25; y', .25; z', .25; v', .25).$$

From Eq. (2), this is equivalent to

$$w_L u(x) + w_{ML} u(y) + w_{MH} u(z') + w_H u(v') > w_L u(x') + w_{ML} u(y') + w_{MH} u(z') + w_H u(v').$$

Because the ranks of z' and v' and the probability distributions are the same on both sides of the inequality, we assume that terms for the common branches can be subtracted from both sides, leaving,

$$w_L u(x) + w_{ML} u(y) > w_L u(x') + w_{ML} u(y').$$

Rearranging, $S' \succ R'$ according to Eq. (2) implies:

$$\frac{w_L}{w_{ML}} > \frac{u(y') - u(y)}{u(x) - u(x')}$$

By repeating the above derivations for different cases, it can be shown that Equation 2 allows violations of branch independence (Birnbaum & McIntosh, 1996). Select outcomes such that $0 < z < v < x' < x < y < y' < z' < v'$, as illustrated in Fig. 2. Gambles containing the small range pair (x, y) will be referred to as S or S' and gambles containing the wide range pair (x', y') will be termed R or R'. Equation (2) implies the following, SR', pattern of violations of branch independence,

$$(x, .25; y, .25; z, .25; v, .25) = S \succ R = (x', .25; y', .25; z, .25; v, .25)$$

and (3a)

$$(x, .25; y, .25; z', .25; v', .25) = S' \prec R' = (x', .25; y', .25; z', .25; v', .25)$$

if and only if

$$\frac{w_{MH}}{w_H} > \frac{u(y') - u(y)}{u(x) - u(x')} > \frac{w_L}{w_{ML}} \tag{3b}$$

The opposite pattern of violations of branch independence, RS',

$$(x, .25; y, .25; z, .25; v, .25) = S \prec R = (x', .25; y', .25; z, .25; v, .25)$$

and (4a)

$$(x, .25; y, .25; z', .25; v' .25) = S' \succ R' = (x', .25; y', .25; z', .25; v', .25)$$

occurs if and only if

$$\frac{w_{MH}}{w_H} < \frac{u(y') - u(y)}{u(x) - u(x')} < \frac{w_L}{w_{ML}} \tag{4b}$$

Expressions (3b) and (4b) show that if the ratios of weights were equal, there would be no violations of branch independence when common outcomes are changed from lowest to highest. Because EU, SEU, and SWU theories assume that weights are independent of rank, these ratios will all be 1 when the probabilities of these distinct outcomes are all equal, and therefore, such nonconfigural theories imply no violations of branch independence in this experiment.

The experimental tactic introduced by Birnbaum and McIntosh (1996) is to vary both the values of the contrast, (x, y) versus (x', y'), and the common outcomes, creating a factorial "fishnet" in which violations implied by Expressions (3b) and (4b) might be "caught." To observe a violation in a finite experiment, it is necessary to use values of the outcomes such that the ratio of differences of utility is "straddled" by the ratios of weights, as in Expressions

(3b) or (4b). Table 1 lists the values of (x, y) and (x', y') used in the present experiment. Four examples of utility functions are also listed in Table 1 to illustrate the impact of $u(x)$ on the predictions.

Cumulative Prospect Model: Predicted Violations of Branch Independence

The cumulative prospect theory (CPT) of Tversky and Kahneman (1992) is a rank- and sign-dependent utility theory, whose representation is the same as that of Luce and Fishburn (1991, 1995). Wakker and Tversky (1993) derived the CPT representation from the assumption of comonotonic independence, whereas Luce and Fishburn (1991, 1995) derived the same representation from a theory of joint receipts. Starmer and Sugden (1989) had noted that the rank-dependent approach of Quiggin (1982) could be generalized and incorporated into the original prospect theory of Kahneman and Tversky (1979), which is what is done in CPT.

In CPT, outcomes are ranked in decreasing value, $x_1 > x_2 > x_3 > \ldots$, so i is the decumulative rank of the outcome; weights of the outcomes are then assumed to follow the expression,

$$w(i) = W(P_i) - W(P_{i-1}), \tag{5a}$$

where $w(i)$ is the weight of outcome, x_i; P_i is the (decumulative) probability that an outcome is greater than or equal to x_i given the gamble, and P_{i-1} is the probability that the outcome is strictly greater than x_i (probability that the outcome is $\geq$ to the next higher value, x_{i-1}). $W(P)$ is a strictly increasing monotonic function such that $W(0) = 0$ and $W(1) = 1$.

TABLE 1

Generic Analysis of Violations of Branch Independence and Distribution Independence

	Contrast S versus R		Utility function			
Row	(x, y)	(x', y')	$u(x) = x$	$u(x) = x^{.88}$	$u(x) = x^{.5}$	$u(x) = \log x$
1	(52, 56)	(11, 97)	1.00	.89	.61	.35
2	(50, 54)	(10, 98)	1.10	.97	.65	.37
3	(45, 49)	(11, 97)	1.41	1.25	.84	.48
4	(40, 44)	(10, 98)	1.80	1.58	1.03	.58
5	(35, 39)	(11, 97)	2.42	2.12	1.39	.79
6	(30, 34)	(12, 96)	3.44	3.01	1.97	1.13

Note. Entries in the last four columns show ratios of differences in utility: $[u(y') - u(y)]/[u(x) - u(x')]$. According to Eq. (2), branch independence will be violated between the cases in which (z, v) are smallest or largest when the ratios of successive weights "straddle" the ratios specified by the experiment and the utility function. For example, if $u(x) = x$, then $S = (z, v, \$40, \$44)$ will be preferred over $R = (z, v, \$10, \$98)$, when $z, v < \$10$ and $S' = (\$10, \$98, z', v')$ will be preferred over $R' = (\$40, \$44, z', v')$ when $z', v' > \$100$ if $w_L/w_{ML} < 1.8 < w_{MH}/w_H$. The inverse-S weighting function assumes that extreme stimuli have greater weight, so $w_L/w_{ML} > 1 > w_{MH}/w_H$, therefore, it implies the opposite pattern of violations of branch independence.

For four-outcome gambles, the inverse-S cumulative prospect model is a special case of Eq. (2) constrained to follow Eq. (5a) and also assumed to follow a particular weighting function. The decumulative weighting function fit by Tversky and Kahneman (1992) in their model is as follows:

$$W(P) = \frac{P^{\gamma}}{[P^{\gamma} + (1 - P)^{\gamma}]^{1/\gamma}} \tag{5b}$$

where γ is the parameter of the weighting function, estimated to be .61. This model implies an inverse-S relationship between $W(P)$ and P, when $\gamma < 1$, as shown by the curve in Fig. 3.

In this paper, we will refer to Eq. (5b) as part of the cumulative prospect *model*, to maintain a distinction between such a weighting function and the more general cumulative prospect *theory*, which allows any strictly increasing $W(P)$ function in Eq. (5a) where $W(0) = 0$ and $W(1) = 1$.

For the general version with any $W(P)$, the highest of four equally likely positive outcomes has a weight, $w_H = W(1/4)$. For the medium-high outcome in this set, $w_{MH} = W(2/4) - W(1/4)$; the weight of the medium-low outcome is $w_{ML} = W(3/4) - W(2/4)$; the lowest outcome has a weight of $w_L = 1 - W(3/4)$.

For the function illustrated in Fig. 3, with $\gamma = .61$, the weights of four equally likely outcomes according to Eqs. (5a) and (5b) (the vertical gaps in Fig. 3), are .43, .15, .13, and .29, for w_L, w_{ML}, w_{MH}, and w_H, respectively, giving ratios of $w_L/w_{ML} = 2.93 > w_{MH}/w_H = .45$. These ratios straddle the ratios of differences in utility in Table 1 for the utility function fit by Tversky and Kahneman (1992), $u(x) = x^{.88}$, in all rows except the last. Therefore, judges should prefer the wide range pair $R = (x', y')$ over the narrow range pair $S = (x, y)$ when the common outcomes are lowest, and they should have the opposite preference

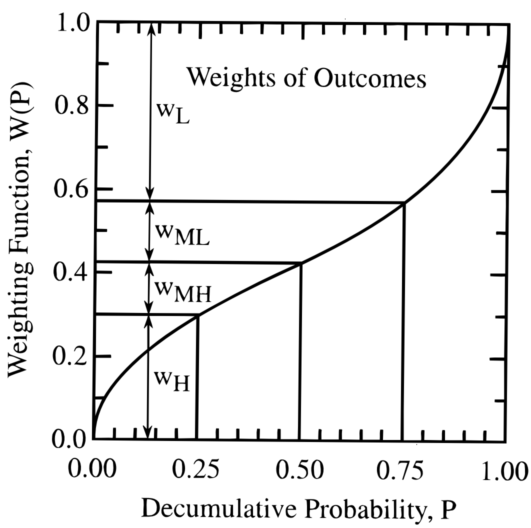

FIG. 3. Weighting function used in the inverse-S model of cumulative prospect theory. Weights are given by vertical differences between values of W(P), where P is the decumulative probability that the outcome is x or higher. Weights of four equally likely outcomes (each with probability = 1/4) have the property that $w_L > w_{ML}$ and $w_H > w_{MH}$; therefore, $w_L/w_{ML} > 1 > w_{MH}/w_H$.

when the common outcomes are the highest (i.e., $S' > R'$), as in Expressions (4a) and (4b), in the first five rows of Table 1. They should prefer $R > S$ and $R' > S'$ in the last row.

Other forms for $W(P)$ have been suggested. Luce and Fishburn (1995) suggested a power function for $W(P)$, which follows from their theory of joint receipts and the assumption that a gamble with more than two outcomes can be thought of as a binary gamble with one outcome as one consequence and a re-normalized gamble on the other outcomes as the other consequence. This implication also follows from any separable theory in which a compound of two independent binary gambles can be represented with the ("rational") product of the probabilities [i.e., suppose $U(x, p, 0) = W(p)u(x)$ and $U\{(x, p, 0)\ q, 0\} \sim U(x, pq, 0)$]. The term *Power-Cumulative Prospect Theory* (P-CPT) refers to the assumptions of Eq. (2), Eq. (5a), and the following weighting function:

$$W(P) = P^{\gamma}, \tag{5c}$$

where γ is the exponent of the power function. This P-CPT requires that the weights of equally likely outcomes should be a monotonic function of their ranks. It can imply SR' or RS' patterns of violation, Expressions (3a–b) or (4a–b), when $\gamma > 1$ or $\gamma < 1$, respectively.

A two-parameter form for $W(P)$ that can describe a family of S and inverse-S shapes can be written as follows:

$$W(P) = \frac{cP^{\gamma}}{cP^{\gamma} + (1 - P)^{\gamma}}, \tag{5d}$$

where c is a constant that reflects the relative weight of higher versus lower-valued outcomes, and γ characterizes the shape of the $W(P)$ curve. When $c = 1$, $W(P)$ has an inverse-S shape for $\gamma < 1$, is linear for $\gamma = 1$, and has an S shape when $\gamma > 1$.

The parameter, c, can be interpreted as an index of risk seeking versus risk aversion. Note that $W(1/2) = c/(c + 1)$ for any γ, so c translates directly into the weight of a higher outcome of probability 1/2. When $c = 1$, $W(1/2) = 1/2$; when $c < 1$ or $c > 1$, $W(1/2) < 1/2$ or $W(1/2) > 1/2$, respectively. [It is important to distinguish the explanation of "risk aversion" as due to weights as opposed to "risk aversion" due to curvature of the utility function. This issue is explored in greater detail in Birnbaum and Sutton (1992) and Birnbaum *et al.* (1992).]

We use the term *Inverse-S Cumulative Prospect Theory* to refer to Eq. (5a), with any weighting function, such as that in Fig. 3, that is "flatter in the middle," such that $w_{ML} < w_L$ and $w_{MH} < w_H$. It follows that, $w_L/w_{ML} > 1$, and $1 > w_{MH}/w_H$; therefore, such a function implies the preference order RS', by Expressions (4a–b). Tversky and Wakker (1995) and Tversky and Fox (1995) used the term "S-shaped" for the function in Fig. 3, which we call "inverse-S"; apparently, they did not consider the possibility that data would soon require a distinction between the "S" and "inverse-S" forms.

Thus, the implications of the particular inverse-S curve of Eq. (5b) also holds for other equations (such as Eq. (5d) with $\gamma < 1$) with these same properties (e.g., Tversky & Fox, 1995; Tversky & Wakker, 1995; Prelec, 1995; Wu & Gonzalez, 1996). The term, Inverse-S Cumulative Prospect Theory (IS-CPT) refers to Eq. (2), Eq. (5a) and the assumptions that $w_{ML} < w_L$ and $w_{MH} < w_H$. The term, S-Cumulative Prospect Theory (S-CPT) refers to Eq. (2), Eq. (5a), and the opposite shape weighting function, such that $w_{ML} > w_L$ and $w_{MH} > w_H$. Expression (5d) with $\gamma > 1$ can follow such an S shape.

Birnbaum and McIntosh (1996) found the opposite pattern from that predicted by IS-CPT. Birnbaum and McIntosh (1996) also showed that the empirical inverse-S relationship between certainty equivalents of two-outcome gambles and probability (the data used by Tversky and Kahneman, 1992, to fit their $W(P)$ curve) can be interpreted as a consequence of a configural weighting function that does not obey Eq. (5a).

With binary gambles, as used by Tversky and Kahneman (1992), Eq. (5b) can be fit if $u(x)$ is known, but Eq. (5a) remains untested. Wu and Gonzalez (1996) confirmed the inverse-S weighting function using a procedure that allows estimation of the weighting function without assuming the form of $u(x)$. However, their procedure also assumes but does not test Eq. (5a).

Thus, the data and theory of Birnbaum and McIntosh (1996) are inconsistent with the weighting function of Tversky and Kahneman (1992); on the other hand, the model of Birnbaum and McIntosh can fit the data of Tversky and Kahneman (1992) as well as the data of Wu and Gonzalez (1996), which had been interpreted as consistent with the inverse-S cumulative weighting function.

Configural Weight Model: Opposite Violations of Branch Independence

Birnbaum and McIntosh (1996) noted that the following model yields predictions that are nearly identical to those of Tversky and Kahneman (1992) for binary gambles ($x > y \geq 0$):

$$U(x, p; y, 1 - p) = w_H u(x) + w_L u(y) \tag{6a}$$

where

$$w_H = \frac{a_H S(p)}{a_H S(p) + a_L S(1 - p)} \tag{6b}$$

and

$$w_L = \frac{a_L S(1 - p)}{a_H S(p) + a_L S(1 - p)} = 1 - w_H \tag{6c}$$

where $a_L = (1 - a_H) = .63$; $u(x) = x$ for $0 \leq x \leq \$150$; and $S(p) = p^{.6}$. Note that Expression 6a is simply a weighted average of the utilities in which the relative weights (in Expression 6b and 6c) sum to 1. The relative weights in Eqs. (6b) and (6c) include a transformation of probability and a configural weight parameter, a_L, (that depends on the judge's point of view, as postulated by Birnbaum and Stegner (1979; see also Birnbaum *et al.*, 1992). For three or more outcomes, the $S(p)$ function is assumed to be invariant, and only the configural weights differ. Each relative weight is always the ratio of the product of a configural weight and $S(p)$, divided by the sum of these products for all outcomes in the gamble.

Equations (6a–c) can be derived from the assumption that judges minimize an asymmetric loss function (Birnbaum *et al.*, 1992, Eq. (6)), allowing also a nonlinear psychophysical function for probability, $S(p)$ and $S(1 - p)$. Note also that Eq. (5d) is a special case of Eq. (6b), where $S(p)$ is a power function. The value of $a_L = .63$ agrees with two other studies of binary gambles: "neutral's" prices for two positive outcomes (Birnbaum *et al.*, 1992) and judged strengths of preference between gambles composed of two equally likely positive outcomes (Birnbaum, Thompson, & Bean, 1997), as well as the median certainty equivalents reported in Tversky and Kahneman (1992).

Birnbaum and McIntosh (1996) estimated configural weight parameters for three equally likely outcomes in the equation, $U(x, y, z) = w_L u(x) + w_M u(y) + w_H u(z)$, where $0 < x < y < z$. The configural weights were .51, .33, and .16 for a_L, a_M, and a_H, respectively, for lowest, middle, and highest, which are the same as the relative weights in the case of equally likely outcomes. These values have ratios close to the ranks of the outcomes (3:2:1 with 1 = highest rank and 3 = lowest), so extrapolating to four outcomes, one might conjecture that relative weights of four equally likely outcomes would be .4, .3, .2, and .1, respectively, proportional to their ranks.

The configural-weight model of Birnbaum and McIntosh (1996) yields predictions for two-outcome gambles that are virtually identical to those of Tversky and Kahneman (1992). It also describes the pattern of results observed by Wu and Gonzalez (1996). However, the configural weight model makes quite different predictions for violations of branch independence and distribution independence in three- and four-outcome gambles. For branch independence, weights with ratios 4:3:2:1 imply Expressions 3a and 3b, opposite the predictions of the IS-CPT. The Birnbaum and McIntosh (1996) model also implies distribution independence, unlike cumulative prospect theory, as shown in the next section.

Distribution Independence

The property of distribution independence that will be tested in this experiment can be defined as follows:

$$S = (x, p; y, q; z, r; v, 1 - p - q - r) \succ R = (x', p; y', q; z, r; v, 1 - p - q - r)$$

$$\text{if and only if} \qquad (7)$$

$$S' = (x, p; y, q; z, r'; v, 1 - p - q - r') \succ R' = (x', p; y', q; z, r'; v, 1 - p - q - r')$$

As distinguished from branch independence, distribution independence asserts that the trade-off between $(x, p; y, q)$ and $(x', p; y', q)$ should be independent of the probability distribution of the common branches (r versus r'), though it would certainly depend on p and q.

Figure 4 illustrates distribution independence. Note that the outcomes on the common branches are the same, but their probabilities differ in the two choices; whereas in branch independence (Fig. 2) the probabilities of the common outcomes are the same, and the outcome values differ. In this paper, we will test distribution independence with equal probabilities, $p = q = .2$, sandwiched between two common components that vary in their probabilities, as depicted in Fig. 4.

For this test, we select outcomes such that $0 < z < x' < x < y < y' < v$, as illustrated in Fig. 4. The gambles compared will be of the form:

$$S = (z, r; x, .2; y, .2; v, .6 - r) \text{ versus } R = (z, r; x', .2; y', .2; v, .6 - r),$$

where S contains the small range pair (x, y), and R contains the large range pair (x', y'); furthermore, $r = .59, .55, .05$, or $.01$. According to the generic rank-dependent model, for example,

$$(z, .55; x, .2; y, .2; v, .05) = S \succ R = (z, .55; x', .2; y', .2; v, .05)$$

$$\text{if and only if} \tag{8a}$$

$$w_L u(z) + w_{ML} u(x) + w_{MH} u(y) + w_H u(v) >$$
$$w_L u(z) + w_{ML} u(x') + w_{MH} u(y') + w_H u(v),$$

where w_L, w_{ML}, w_{MH}, and w_H, are the relative, configural weights of the Lowest, Medium-Low, Medium-High, and Highest outcomes at probabilities of .55, .2, .2, and .05, respectively (note that these weights will not have the same values as in Eq. (2) with equal probabilities). Because the distributions of probability

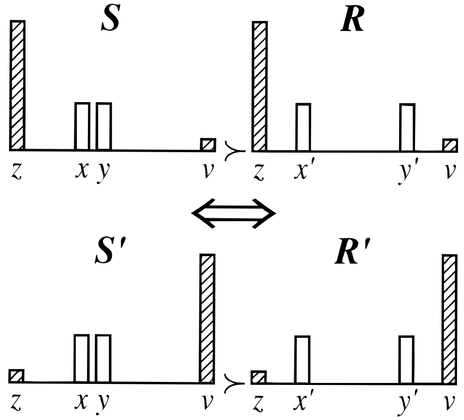

FIG. 4. A test of distribution independence: S is preferred to R if and only if S' is preferred to R'. Bars represent probabilities of outcomes; note that probabilities of common outcomes (z, v) are the same within each comparison but are different between comparisons of S versus R and S' versus R'. Stripping away the common outcomes (shown with shaded probabilities) would leave the same contrast in both cases, between (x, y) and (x', y').

and the ranks of common outcomes are the same on both sides, it is assumed that we can subtract the products for the common branches, leaving,

$$w_{ML}u(x) + w_{MH}u(y) > w_{ML}u(x') + w_{MH}u(y')$$

which implies

$$\frac{w_{ML}}{w_{MH}} > \frac{u(y') - u(y)}{u(x) - u(x')}. \tag{8b}$$

Expression (8b) indicates that if the ratio of weights of the two components (whose probabilities are fixed) is independent of the distribution, and if the utilities are also independent of the distribution, then the preference will be independent of the distribution. However, if w_{ML} and w_{MH} depend on the configuration of probabilities in the common branches (r or r'), then the ratio on the left could change, which can change the direction of preference for some values of (x, y) and (x', y').

Cumulative Prospect Theory Implies Violations of Distribution Independence

The weighting function of Tversky and Kahneman (1992) implies changes in weights as a function of the distribution of probabilities on the common branches, as illustrated in Figs. 5 and 6; therefore, it predicts violations of distribution independence. Figure 5 shows that when the lowest outcome has a probability of .55, (as in the S versus R comparison in Fig. 4), the medium-high outcome will have more weight than the medium low, $w_{MH} > w_{ML}$; however,

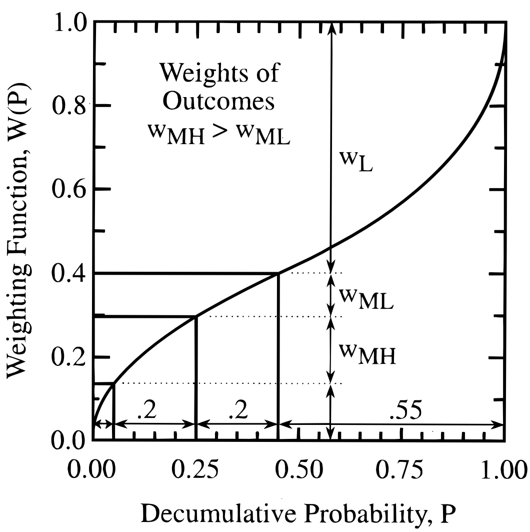

FIG. 5. According to inverse-S weighting function of CPT, the weights of two equally-probable, intermediate outcomes depend on the distribution of extreme outcomes. This figure illustrates the weights for gambles of the form (z, .55; x, .2; y, .2; v, .05), including S and R of Fig. 4; in this case, $w_{MH} > w_{ML}$. The same relation holds for gambles of the form (z, .59; x, .2; y, .2; v, .01).

when the lowest outcome has a probability of .05, (as in the S' versus R' comparison in Fig. 4), Figure 6 illustrates how the weights change rank order, $w_{MH} < w_{ML}$. Changes in preference order (violations of distribution independence) are predicted whenever the change in the ratio of weights in Expression 8b causes them to "straddle" the ratio of differences in utility, as in Table 1.

For the particular parameters given in Tversky and Kahneman ($\gamma = .61$, $u(x) = x^{.88}$), the weights in Fig. 5 are .605, .105, .159, and .132 for w_L, w_{ML}, w_{MH}, and w_H, respectively, corresponding to probabilities of .55, .2, .2, .05. In Fig. 6, the weights are .207, .225, .121, and .447 for probabilities of .05, .2, .2, .55, for w_L, w_{ML}, w_{MH}, and w_H, respectively. Thus the ratio, w_{ML}/w_{MH}, changes from .66 in Fig. 5 to 1.86 in Fig. 6. These ratios straddle the ratios of differences in utility in the first four rows of Table 1, assuming $u(x) = x^{.88}$. The *inverse-S CPT* model therefore implies that judges should prefer the riskier, wide range gamble (R) when the lowest outcome has the greatest probability (either .55 or .59), and the judge should prefer the safer, small range gamble (S') when the lowest outcome has a small probability (when the lowest outcome has probability .05 or .01), in the first four rows of Table 1. In other words, IS-CPT implies the RS' pattern of violations of distribution independence, as follows:

$$(z, .55; x, .2; y, .2; v, .05) = S < R = (z, .55; x', .2; y', .2; v, .05)$$

$$\text{and} \qquad (9)$$

$$(z, .05; x, .2; y, .2; v, .55) = S' > R' = (z, .05; x', .2; y', .2; v, .55)$$

Thus, the cumulative assumption (Eq. (5a)) implies that the preference order

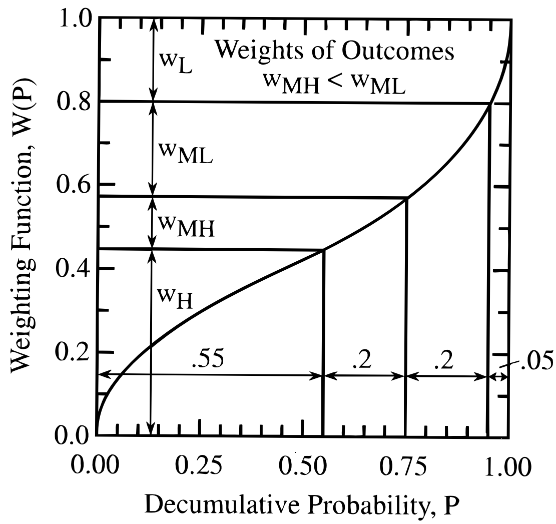

FIG. 6. According to the inverse-S weighting function, the weights of the intermediate outcomes will have the opposite order from Fig. 5, $w_{MH} < w_{ML}$, in gambles of the form (z, .05; x, .2; y, .2; v, .55). The same relation is implied for gambles of the form (z, .01; x, .2; y, .2; v, .59). Figs. 5 and 6 show that the inverse-S weighting function implies violations of the pattern, $R > S$ and $S' > R$, for tests such as that illustrated in Fig. 4.

can be changed by changing the position of fixed outcomes in the cumulative distribution, which is done by changing the probability distribution of the common outcomes. The opposite pattern of violations would be implied by Eq. (5a), if the $W(P)$ function followed an S-shape instead of the inverse-S function.

If judges edited the gambles being compared and cancelled common components (Tversky, 1972; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), then there would be no systematic violations of either branch independence or distribution independence. Note that in either Fig. 2 or Fig. 4, "trimming" the common branches between S and R and between S' and R' leaves the same comparison.

Birnbaum and McIntosh Model: No Violations of Distribution Independence

The configural weight model of Birnbaum and McIntosh (1996) implies no violations of distribution independence (Birnbaum, 1997). [The "revised" configural weight model of Birnbaum and Stegner (1979, Eq. 10) violates distribution independence. This older model will be taken up in the next section.] In the Birnbaum and McIntosh (1996) model, absolute weights depend entirely on the ranks of the outcomes and their probabilities. Although relative weights depend on the probability distribution (affecting strengths of preference), ratios of relative weights (as in Expression(8b)) will be independent of the probability distribution. Therefore, the direction of preference should not be affected by the distribution of probabilities of the common outcomes.

Distribution independence also holds for a more general configural weight model in which the weight of an outcome is any positive-valued function of the outcome's probability and the rank of the outcome, $w(i) = f(p_i, k_i)$, where f is the function, p_i is the probability and k_i is the rank of outcome x_i. In this model, the rank of an outcome (k_i) does not refer to its cumulative distribution value, but instead to the rank of the outcome's value among the distinct values within the gamble, independent of their probabilities. For four-outcome gambles, these would be rank = 1, for the lowest outcome, up to rank = 4 for the highest. Suppose,

$$S = (z, r; x, .2; y, .2; v, .6 - r) \succ R = (z, r; x', .2; y', .2; v, .6 - r) \quad (10a)$$

This preference relation holds according to this configural weight model if and only if

$$\frac{f(r, 1)u(z) + f(.2, 2)u(x) + f(.2, 3)u(y) + f(.6 - r, 4)u(v)}{f(r, 1) + f(.2, 2) + f(.2, 3) + f(.6 - r, 4)} > \frac{f(r, 1)u(z) + f(.2, 2)u(x') + f(.2, 3)u(y') + f(.6 - r, 4)u(v)}{f(r, 1) + f(.2, 2) + f(.2, 3) + f(.6 - r, 4)} \quad (10b)$$

because the gambles have the same probability distribution, the denominators on both sides are the same, and it is assumed that we can multiply both sides by this common factor. Next, because the common terms hold the same ranks in both gambles, we can subtract the common terms, $f(r, 1)u(z)$ and $f(.6 - r,$

4)$u(v)$, from both sides. Then, add new common terms to both sides, $f(r', 1)u(z)$ and $f(.6 - r', 4)u(v)$. We can divide both sides by the same factor, $f(r', 1) + f(.2, 2) + f(.2, 3) + f(.6 - r', 4)$. The result then holds if and only if

$$S' = (z, r'; x, .2; y, .2; v, .6 - r') \succ R' = (z, r'; x', .2; y', .2; v, .6 - r'), \tag{10c}$$

which shows that this configural weight model implies distribution independence. The model of Birnbaum and McIntosh (1996) is a special case of this model in which the function, f, is the product of a function of probability and a configural-weight parameter that depends on the outcome's rank and point of view, i.e., $f(p_i, k_i) = a_V(k_i)S(p)$. Point of view is assumed to be fixed within a given task and instructional set. Therefore, the Birnbaum and McIntosh (1996) model implies distribution independence in this experiment.

Original prospect theory (Kahneman & Tversky, 1979) was not explicated for the case of four-outcome gambles, but its representation can be extended along the lines suggested in Tversky and Kahneman (1986) to make an implication for the present experiment. The original theory differs from cumulative prospect theory in that weights of outcomes are presumed to be a function of the probabilities of the outcomes rather than computed from the functional of Expression (5a). Extending the original prospect model to four-outcome gambles (by assuming utility of a gamble to be the sum of weighted products of a function of probability of the outcomes and a function of the outcomes), this extension of original prospect theory would also imply no violations of distribution independence, because the common terms can be subtracted off both sides and replaced by new common terms. Tversky and Kahneman (1986) also proposed that judges will edit and cancel common terms, and this principle provides a separate argument that also implies distribution independence. EU and SWU theories also imply distribution independence.

Birnbaum and Stegner (1979) Model: Violations of Distribution Independence

Birnbaum and Stegner (1979, Eq. 10) proposed a "revised" configural weight model in which configural weights are transferred in proportion to the absolute weight of the stimulus that loses weight. Consider gambles of the form, $G = (x_1, p_1; x_2, p_2; \ldots x_j, p_j; \ldots; x_i, p_i; \ldots; x_n, p_n)$, where the outcomes are ordered such that $x_1 < x_2 < \cdots < x_j < \cdots x_i < \cdots < x_n$, and $\Sigma\, p_j = 1$. This model can be written as follows:

$$U(G) = \frac{\sum_{i=1}^{n} S(p_i)u(x_i) + \sum_{i=2}^{n}\sum_{j=1}^{i-1} (u(x_i) - u(x_j))\omega(i, j, G)}{\sum_{i=1}^{n} S(p_i)} \tag{11}$$

where $U(G)$ is the utility of the gamble, $S(p)$ is a function of probability, $u(x)$

is the utility function of money, and $\omega(i, j, G)$ is the configural transfer of weight between outcomes x_i and x_j. Note that $x_i > x_j$, so if the configural term is negative, then the higher-valued outcome loses weight and the lower valued outcome gains this same weight.

Birnbaum and Stegner (1979) compared two versions of the above model, one in which the configural weight transfer is independent of the sources' absolute weights, and another in which the weight transferred is proportional to the absolute weight of the stimulus losing weight (Birnbaum & Stegner, Eq. 10). These different versions of configural weighting are not distinguishable for an experiment such as Birnbaum and McIntosh (1996), where the outcomes are equally likely, but they can be differentiated in the present experiment, which varies the probabilities of the outcomes.

The latter model, where configural weight transferred is proportional to the absolute weight of the stimulus losing weight, can be simplified by the following assumption,

$$\omega(i, j, G) = S(p_i)w(j, n), \text{ if } w(j, n) \leq 0 \tag{12a}$$

$$\omega(i, j, G) = S(p_j)w(j, n), \text{ if } w(j, n) > 0, \tag{12b}$$

where $w(j, n)$ is a function of the rank of the item gaining weight (j) and the number of distinct outcomes (n) in gamble, G; and $w(j, n) < 0$. Simplifying further, if $w(j, n) = -1/(n + 1)$, then the weights of two, three, and four equally likely outcomes would be $(\frac{2}{3}, \frac{1}{3})$, $(\frac{3}{6}, \frac{2}{6}, \frac{1}{6})$, and (.4, .3, .2, .1), for lowest to highest outcomes, respectively, the same pattern as postulated above for the Birnbaum and McIntosh (1996) model. Equations (11) and (12) can be distinguished from that model in the tests for violations of distribution independence, however.

For example, if $S(p) = p$, $u(x) = x$, and letting $w(j, n) = -1/(n + 1)$, Equations (11) and (12) imply the choice pattern SR' for a test of distribution independence in Row 3 of Table 1, as follows:

$$S = (\$2, .55; \$45, .2; \$49, .2; \$108, .05) > R = (\$2, .55; \$10, .2; \$98, .2; \$108, .05)$$

and

$$S' = (\$2, .05; \$45, .2; \$49, .2; \$108, .55) < R' = (\$2, .05; \$10, .2; \$98, .2; \$108, .55).$$

The values of utility in this case are $U(S) = 13.8 > U(R) = 13.0$ and $U(S') = 53.5 < U(R') = 54.9$. In this example, the weights of the outcomes, w_L, w_{ML}, w_{MH}, and w_H, would be .64, .21, .13, and .01 for probabilities of .55, .2, .2, and .05; and .24, .31, .27, and .22 for .05, .2, .2, and .55, respectively. Note that $w_{ML}/w_{MH} = 1.61$ in the first distribution (S versus R), and 1.15 in the second (S' versus R'), straddling the value in Table 1 for the third row, assuming $u(x) = x$. Thus, this model implies the SR' pattern with parameters that also account for the SR' pattern of violations of branch independence.

This model also violates asymptotic independence (Birnbaum, 1997), a property that also distinguishes it from the differential weight averaging model (Birnbaum, 1973; Riskey & Birnbaum, 1974).

Summary of Predictions

Table 2 Summarizes the predictions of the various models. Except for EU and SWU theories and the theory that judges edit and eliminate common components from gambles before comparison, all of the theories considered here can account for violations of branch independence. However, different weighting functions imply different patterns of violations. The inverse-S weighting function of CPT implies the opposite pattern from that predicted by the parameters and model of Birnbaum and McIntosh (1996) or comparable parameters in the Birnbaum and Stegner (1979) model.

Distribution independence is implied by the Birnbaum and McIntosh (1996) model and the extension of original prospect theory (Kahneman & Tversky, 1979). However, distribution independence should be violated in a fashion predicted by the cumulative weighting function according to cumulative prospect models. According to the cumulative prospect models, the pattern of violations of distribution independence follow from the cumulative weighting function and should be compatible with the pattern of violations of branch independence. A similar connection is implied in the configural weight model of Birnbaum and Stegner (1979).

METHOD

Instructions

Judges received printed instructions, which were also read aloud to them. Instructions read (in part) as follows:

TABLE 2

Summary of Predictions of Models

	Property tested	
Model	Branch independence	Distribution independence
EU, SWU	No violations	No violations
Editing	No violations	No violations
OPT without editing	Violations (?)	No violations
Birnbaum–McIntosh	SR' violations	No violations
IS-CPT	RS' violations	RS' violations
P-CPT	Violations depend on γ	Violations depend on γ
S-CPT	SR' violations	SR' violations
Birnbaum–Stegner	SR' violations	SR' violations

Note. EU and SWU refer to Expected Utility theory and Subjectively Weighted Utility theory, respectively. Editing refers to the theory that judges edit and eliminate common components prior to comparison (Kahneman & Tversky, 1979). Original Prospect Theory (OPT), extended to four outcomes, implies distribution independence; it is unclear how extended OPT violates branch independence. Birnbaum–McIntosh refers to Birnbaum and McIntosh (1996) configural weight model. IS-CPT refers to Inverse-S cumulative prospect theory; S-CPT refers to Cumulative Prospect theory using an S-shaped $W(P)$ function; P-CPT refers to cumulative prospect theory with a power function for $W(P)$; Birnbaum-Stegner refers to the "revised" configural weight model of Birnbaum and Stegner (1979).

. . . On each trial, you will be offered a comparison between two gambles. Your task is to decide which of the two gambles you would prefer to play and to judge how much you would pay to play your preferred gamble rather than the other gamble. . . .

Choices were presented as in the following example:

$$302. \_\_\_\_ \quad \frac{.34 \;\; .33 \;\; .33}{\$5 \; \$25 \; \$100} \qquad \frac{.20 \; .25 \; .55}{\$20 \; \$40 \; \$60}$$

Would you prefer the gamble on the left (34 chances out of 100 to get $5, 33 chances to get $25, and 33 chances to get $100) or the gamble on the right (20 chances to get $20, 25 chances to get $40, and 55 chances to get $60)? . . . .

For each gamble, you can think of a can containing 100 identical slips of paper with different amounts written on them. Since the slips are equally likely and one will be chosen at random, the probability of each outcome is the number of slips with that outcome, divided by 100. Each trial displays the probabilities and values of all possible outcomes for each gamble. Probabilities will always sum to 1 in each gamble.

Judges circled the gamble they would prefer to play and then judged the amount they would be willing to pay to receive their preferred gamble rather than the other gamble. For purposes of data analysis, a negative sign was associated with choice of the gamble on the left.

Designs

Branch independence design. The first subdesign consisted of 12 comparisons of the form (z, .25; v, .25; x, .25; y, .25) versus (z, .25; v, .25; x', .25; y', .25), constructed from a 6 by 2, (x, y) versus (x', y') by (z, v) or (z', v') factorial design, as illustrated in Fig. 2. The six levels of (x, y) versus (x', y') were as listed in Table 1: ($52, $56) versus ($11, $97), ($50, $54) versus ($10, $98), ($45, $49) versus ($11, $97), ($40, $44) versus ($10, $98), ($35, $39) versus ($11, $97), ($30, $34) versus ($12, $96), and the 2 levels of (z, v) or (z', v') for the six contrasts were ($2, $3) or ($108, $113), ($3, $4) or ($109, $112), ($2, $5) or ($110, $116), ($2, $3) or ($108, $119), ($4, $6) or ($107, $113), and ($5, $6) or ($111, $118), respectively.

Subdesign 2, testing branch independence with three-outcome gambles, consisted of 12 comparisons of the form (z, .5; x, .25; y, .25) versus (z, .5; x', .25; y', .25); constructed from a 6 by 2 (x, y) versus (x', y') by (z or z') factorial design. The six levels of (x, y) versus (x', y') were the same as in Subdesign 1, and the 2 levels of z(z or z') for the six contrasts were ($2 or $108), ($4 or $107), ($3 or $109), ($5 or $111), ($6 or $113), ($4 or $108), respectively.

Distribution independence design. Subdesigns 3 and 4 consisted of 24 choices of two forms, $S = (z, r; x, .2; y, .2; v, .6 - r)$ versus $R = (z, r; x', .2; y', .2; v, .6 - r)$, and $S' = (z, .6 - r; x, .2; y, .2; v, r)$ versus $R' = (z, .6 - r; x', .2; y', .2; v, r)$, constructed from a, 6 by 2 by 2, (x, y) versus (x', y') by (z, v) by comparison (S versus R or S' versus R'), factorial design, where $r = .59$ or .55 in Subdesigns 3 and 4, respectively. The six levels of contrast, (x, y) versus (x', y'), were the same as in the branch independence designs, and the values of (z, v) for the contrasts were ($2, $108), ($3, $109), ($4, $110), ($5, $108), ($2, $107), ($3, $111), respectively.

Filler design. There were 60 "filler" trials with choices between two-outcome gambles in which there were no common branches, outcomes, or probabilities. These trials were of the form (x, q; y, $1 - q$) versus (x', p; y', $1 - p$). Subdesign 5 consisted of 36 comparisons of the form ($55, .25; $59, .75) versus ($12, $1 - p$; $96, p), ($50, .35; $54, .65) versus ($10, $1 - p$; $98, p), ($45, .45; $49, .55) versus ($11, $1 - p$; $97, p), ($40, .55; $44, .45) versus ($10, $1 - p$; $98, p), ($35, .65; $39, .35) versus ($11, $1 - p$; $97, p), or ($30, .75; $34, .25) versus ($12, $1 - p$; $96, p), constructed from a 6 by 6, Comparison by (p), factorial design. The six levels of p for each comparison were .20, .30, .40, .50, .60, and .70. Subdesign 6 included 18 comparisons of the form ($75, .25; $79, .75) versus ($12, $1 - p$; $96, p), ($70, .35; $74, .65) versus ($10, $1 - p$; $98, p), ($65, .45; $69, .55) versus ($11, $1 - p$; $97, p), ($60, .55; $64, .45) versus ($10, $1 - p$; $98, p), ($55, .25; $59, .75) versus ($11, $1 - p$; $97, p), or ($50, .35; $54, .65) versus ($12, $1 - p$; $96, p), constructed from a 6 by 3 (x, y) versus (x', y') by (p) design. The three levels of p for each comparison were .80, .90, and .95. Subdesign 7 included 6 comparisons as follows: ($75, .25; $79, .75) versus ($0, .05; $96, .95), ($70, .35; $74, .65) versus ($0, .05; $96, .95), ($65, .45; $69, .55) versus ($0, .05; $96, .95), ($60, .55; $64, .45) versus ($0, .05; $96, .95), ($55, .25; $59, .75) versus ($0, .05; $96, .95), ($50, .35; $54, .65) versus ($0, .05; $96, .95). There were 10 additional "filler" choices between three-outcome gambles, using probabilities of (.33, .34, .33) and (.30, .40, .30).

Check design. The "check" design consisted of 12 comparisons with transparent dominance, in which both gambles were the same, except one outcome of one gamble was higher, or the probability of a higher outcome was greater in one gamble. There were 3 comparisons of the form (x, .25; y, .25; $50, .5) versus ($x$, .25; y; .25; $90, .5), in which the 3 levels of (x, y) were ($2, $4), ($2, $108), ($108, $111); 3 comparisons of the form: (x, .1; y, .1; $92, .8) versus ($x$, .1; y; .1; $42, .8), in which the 3 levels of (x, y) were ($3, $5), ($3, $109), ($109, $112); 3 trials as follows: ($5, .6; $50, .2; $90, .2) versus ($5, .2; $50, .6; $90, .2), ($5, .2; $50, .6; $90, .2) versus ($5, .2; $50, .2; $90, .6) and ($5, .6; $50, .2; $90, .2) versus ($5, .2; $50, .2; $90, .6); 3 trials as follows: ($4, .1; $40, .8; $93, .1) versus ($4, .8; $40, .1; $93, .1), ($4, .1; $40, .1; $93, .8) versus ($4, .1; $40, .8; $93, .1), and ($4, .1; $40, .1; $93, .8) versus ($4, .8; $40, .1; $93, .1). Half of these required the judge to choose the gamble on the left, and half on the right.

Procedure

The 130 choices from all eight subdesigns were intermixed and printed in booklets in pseudo random order with the restrictions that successive trials did not repeat outcome values, and no two successive trials came from the same subdesign. Each booklet contained 3 pages of instructions with example trials, 10 warm-up trials, followed by 5 unlabeled practice trials and 130 experimental trials.

The experimenter checked the responses to the first 10 warm-up trials. Initial examples were very simple, such as the choices in the "check" design, and when judges violated transparent dominance in such trials, the experimenter would ask the judge to explain the choice, directing the judge to reread the instructions as needed. When the warm-ups satisfied transparent dominance, judges were directed to proceed to the experimental trials.

Judges completed the experiment within 1 h, working at their own paces.

Research Participants

The judges were 100 undergraduates enrolled in Introductory Psychology who made one or fewer violations of transparent dominance on the 12 "check" trials (27 of these had one violation). Additional judges were tested who violated transparent dominance two or more times and were excluded from the analysis.

RESULTS

Violations of Branch Independence

Table 3 presents the number of judges who exhibited each pattern of preferences in the branch independence designs. Judges could respond in each of four ways for each combination of outcomes, shown as rows in the table. SS' indicates that judges preferred the gamble containing the small range, $S = (x, y)$, outcomes over the gamble containing the wider range outcomes, $R = (x', y')$, for both values of z, v. RR' indicates $R \succ S$ and $R' \succ S'$. SR' indicates

TABLE 3

Tests of Branch Independence with Four- and Three-Outcome Gambles

		Gamble type							
		Four outcomes (z, .25; v, .25; x, .25; y, .25)				Three outcomes (z, .5; x, .25; y, .25)			
S (x, y)	R (x', y')	SS'	SR'	RS'	RR'	SS'	SR'	RS'	RR'
($52, $56)	($11, $97)	33	**34***	10	23	45	**27***	6	22
($50, $54)	($10, $98)	37	**29***	10	24	44	**22***	5	29
($45, $49)	($11, $97)	26	**30***	13	31	37	**24**	15	24
($40, $44)	($10, $98)	22	**28**	16	34	31	**29***	7	33
($35, $39)	($11, $97)	11	**21**	18	50	22	**27***	7	44
($30, $34)	($12, $96)	13	**28***	11	48	14	**32***	11	43

Note. Each entry is the number (of 100) judges with each preference pattern in Subdesigns 1 and 2. SS′ indicates preference for the smaller range, $S = (z, .25; v, .25; x, .25; y, .25) \succ R = (z, .25; v, .25; x', .25; y', .25)$ when $z, v <$ \$10 and $S' \succ R'$ when $z', v' >$ \$100; SR' indicates the preference pattern of Expression 3a, i.e., $S \succ R$ (when $z, v <$ \$10) and $R' \succ S'$ (when $z', v' >$ \$100); RS' indicates the opposite $R \succ S$ and $S' \succ R'$, predicted by IS-CPT (Expression 4a); and RR' indicates preference for R and R' in both cases. Asterisks indicate cases in which the violations of branch independence are significantly different; in all cases, they are more numerous for SR' (shown in bold type) than for RS'.

the pattern of violations predicted by Expression (3a) namely $S > R$ when the common outcome(s) are lowest (z, $v < \$10$) and $R' > S'$ when the common outcome(s) are highest (z', $v' > \$100$). RS' indicates the opposite pattern, predicted by inverse-S cumulative prospect theory (Expression (4a)). The first four columns of data show the results for gambles composed of four equally likely outcomes. The last four columns of data show the results for three-outcome gambles of Subdesign 2, in which $r = .50$.

If the data in Table 3 were perfectly consistent with branch independence, the entries under SR' and RS' would all be zero. If violations of branch independence were due to random errors, frequencies of SR' and RS' should not differ systematically from each other. Instead, in all six rows of both designs, the pattern SR' (shown in bold type) is more frequent than the pattern RS'. The probability of obtaining 12 contrasts in the same direction by chance, if violations of branch independence were unsystematic, is $(1/2)^{12} = .0002$, which is significant (the term, "significant" refers to $\alpha = .05$ throughout). The two types of violations, SR' versus RS', were also tested separately in each row against the null hypothesis that their frequencies resulted from a binomial process with $p = .5$; asterisks in Table 3 show that 9 of 12 cases deviate significantly from the null hypothesis.

The predominant pattern of violations is consistent with Expression (3a) and the pattern observed in previous choice experiments with three-outcome gambles (Birnbaum & McIntosh, 1996, submitted; Weber & Kirsner, 1997).

The pattern of violations is opposite that implied by the inverse-S weighting function of the cumulative prospect model of Tversky and Kahneman (1992). Because medium outcomes are supposed to have lower weight in that model, judges should have had the RS' pattern in the first four rows of Table 3, contrary to the data.

A similar pattern of violations was observed in the mean and median strength of preference judgments of Subdesigns 1 and 2. (Note that Table 3 only analyzes the choice relation, and not the judgment of strength of preference). Analysis of variance of strength of preference judgments showed significant effects of the common outcomes in four-outcome gambles, $F(1, 99) = 7.08$, and in three-outcome gambles, $F(1, 99) = 5.84$. On the average, judges offered to pay \$1.18 for the lower range, four-outcome gamble when z, $v < \$10$ (i.e., to receive S instead of R), but offer to pay \$3.79 for the wide range pair when z, $v > \$100$ (i.e., to receive R' instead of S'). For three-outcome gambles, the corresponding values were \$2.54 to buy the narrow range pair when z, $v < \$10$, and \$1.33 for the wide range pair, when z, $v > \$100$. Effect of rows was also significant in both analyses, $F(5, 495) = 11.91$ and 13.97.

Violations of Distribution Independence

Table 4 shows the percentage of choices in each pattern for the distribution independence design, with a separate set of four columns for each distribution of probabilities of the extreme outcomes. According to distribution independence, preferences should not change as a function of the probabilities of the common

TABLE 4

Tests of Distribution Independence

		(z, .59; x, .2; y, .2; v, .01)				(z, .55; x, .2; y, .2; v, .05)			
S (x, y)	R (x′, y′)	SS'	SR'	RS'	RR'	SS'	SR'	RS'	RR'
($52, $56)	($11, $97)	54	**20**	10	16	47	**22**	14	17
($50, $54)	($10, $98)	49	**14**	11	26	47	**19***	7	27
($45, $49)	($11, $97)	43	**23***	6	28	41	**19**	10	30
($40, $44)	($10, $98)	37	**21**	11	31	38	**22***	10	30
($35, $39)	($11, $97)	34	**15**	10	41	31	**23***	6	40
($30, $34)	($12, $96)	23	**30***	5	42	26	**26***	8	40

Note. Each entry is the number of judges who exhibited each preference pattern in Subdesigns 3 and 4. S indicates preference for $S = (z, r; x, .20; y, .20; v, .6 - r)$ over $R = (z, r; x', .20; y', .20; v, .6 - r)$ when $r = .59$ and S' indicates preference for $S' = (z, r'; x, .20; y, .20; v, .6 - r')$ over $R' = (z, r'; x', .20; y', .20; v, .6 - r')$ when $r' = .01$; RS' indicates the preference pattern predicted by IS-CPT, $R > S$ and $S' > R'$; SR' indicates the opposite preference pattern, $S > R$ and $R' > S'$; and RR' indicates consistent preference for the wide range gamble in both distributions. The last four columns involve choices in which $r = .55$ and $r' = .05$. Asterisks indicate significant asymmetry in violations. In all choices in both subdesigns, there are more SR' (bold type) violations than RS', opposite the preidctions of IS-CPT.

outcomes, so except for error, all choices should be either SS' or RR'. If the violations, SR' and RS', were unsystematic, then they should split equally in either direction. Instead, in all 12 cases (all six in the first four columns and all six cases in the second four columns, shown in bold type), SR' is more frequent than RS'. This pattern is opposite the pattern predicted by the inverse-S weighting function of the cumulative prospect model (Expression (9)). Systematic violations are also inconsistent with the configural weight model of Birnbaum and McIntosh (1996).

Analysis of variance of the strength of preference judgments corresponding to Table 4 showed significant effects of the distribution, $F(1, 99) = 14.33$ and $F(1, 99) = 15.66$, for $r = .59$ or .01 and for $r = .55$ or .05, respectively. The effects of rows were significant in both cases, and the interactions were not.

Analyses of Individual Data

If individual judges have different utility functions but the same weights, then they would be expected to violate branch independence in different rows of Table 1. For example, if $u(x) = x$, and if a judge had the pattern of weights estimated from the group data of Birnbaum and McIntosh (with ratios of 3:2:1), then that judge should show exactly one SR' violation of branch independence in the three-outcome test in Table 3, in the fourth row. The first three rows should be SS' and the last two should show RR'. Another judge with the same weights and $u(x) = x^{.5}$ would show one SR' violation in Row 6 of the design, with the first five rows showing SS'. Similarly, if judges have different weights, they could also show different patterns. As a further complication, empirical choices may contain variability due to error.

For each judge in each subdesign, we calculated the difference between the frequencies of the two patterns of violations of branch independence and distribution independence, SR' versus RS', summing over rows in the design. For Subdesign 1 (four-outcome gambles), there were 51 judges with more SR' violations, 24 with more RS' violations, and 25 with no difference, including 9 who never violated branch independence (in all 9 of these cases the judge chose S or R in every row). For three-outcome gambles (Subdesign 2), there were 58 judges who showed more violations of the type SR' against 16 who showed more RS' violations, with 26 judges having no difference (among these 26, there were 20 who satisfied distribution independence in this design, of whom 19 consistently chose either SS' or RR' on all six choices).

In Subdesign 3, testing distribution independence ($r = .59$, $r' = .01$), testing violations of distribution independence, there were 55 judges showing more SR' violations, only 14 showing the pattern predicted by inverse-S function (RS'), and 31 who showed no difference, including 20 with no violations, of whom 19 consistently chose R or S in all six rows. In Subdesign 4 ($r = .55$, $r' = .05$), there were 49 judges who had more SR' violations, 14 with more RS' violations, and 37 with no difference, including 20 who never violated distribution independence, of whom 19 consistently chose R or S in all rows.

It is interesting that among those judges who changed preferences between rows [as (x, y) is decreased], no more than one judge consistently satisfied branch independence or distribution independence in any subdesign; every other judge who changed preferences also showed at least one violation.

The counts of individual data reinforce the group analyses, showing that the pattern SR' is characteristic of more judges in all four subdesigns than the opposite pattern, RS'.

These four contrasts for each judge should be interrelated if they are due to real individual differences, rather than to error. For example, if the judge has more violations of SR' in Subdesign 1 with four-outcome gambles, the same judge should also have more SR' in Subdesign 2, with three-outcome gambles, if the violations are due to the judge's weighting pattern rather than chance. (Note that comparison of these subdesigns also provides an indirect test of a coalescing equivalence that assumes that when two outcomes are equal, the two outcomes can be coalesced by adding their probabilities.)

We cross-tabulated the patterns of violation for all six pairs of the Subdesigns 1–4. These six cross-tabulations are shown as rows of Table 5.

The columns of Table 5 show the nine possible combinations of patterns of violations (more RS', = equal split or none, more SR') in two subdesigns. The entry of 44 in the first row, last column shows that of the 51 judges who showed more SR' violations in Subdesign 1, 44 (86%) also showed more SR' violations in Subdesign 2. The entry of 8 in the first row, first column shows by contrast that of the 24 judges who showed more RS' violations in Subdesign 1, 8 (33%) also showed more RS' violations in Subdesign 2. Similar results were observed in the cross-tabulations between other pairs of designs. Asterisks indicate significant differences; in all cases, there are more judges showing the pattern of Expression(3a) (SR') in two subdesigns (rightmost column, in bold type)

TABLE 5

Cross-Tabulation of Violations of Branch Independence and Distribution Independence between Subdesigns

	Combination of modal preferences								
Designs	RS'&RS'	RS'&=	RS'&SR'	=&RS'	=&=	=&SR'	SR'&RS'	SR'&=	SR'&SR'
1 by 2	8	10	6	6	11	8	2	5	**44***
1 by 3	6	8	10	0	15	10	8	8	**35***
1 by 4	7	9	8	2	13	10	5	15	**31***
2 by 3	4	4	8	2	12	12	8	15	**35***
2 by 4	6	4	6	2	16	8	6	17	**35***
3 by 4	4	6	4	6	12	13	4	19	**32***

Note. Each judge is classified by the more frequent pattern of violations of branch independence and distribution independence over rows. For example, the 6 in the first row under the column labeled RS' & SR', indicates that six judges showed more RS' violations in Subdesign 1 and more SR' violations in Subdesign 2. "=" indicates that the judge showed an equal split of violations (including 0 violations).

than showing the pattern of Expression (4a) (RS') in two subdesigns (leftmost column).

According to cumulative prospect theory (Expression (5a)), violations in Subdesigns 1 and 2 should be correlated with violations in Subdesigns 3 and 4. Although the predominant pattern of the data is opposite the pattern predicted by inverse-S weighting functions (including Expression (5b)), the data are consistent with the connection implied by Expression (5a). The two types of violations appear related in the manner that would be consistent with an S-shaped cumulative weighting function for Expression (5a), rather than an inverse-S. The pattern is also consistent with the "revised" configural weight model (Expressions (11) and (12)).

TABLE 6

Median Parameter Estimates and Indices of Fit of CPT Models

	Parameters					Indices of Fit	
Model	γ	β	c	a	b	SUM	$-\log(\Pi P)$
S-CPT (5)	1.59	.82	.31	0.26	1.30	17495	52.0
S-CPT (4)	1.41	(1.0)	.29	0.08	0.49	19514	52.7
P-CPT (4) $u(x) = x^{\beta}$	1.84	1.07	—	0.06	0.41	19305	53.4
P-CPT (4) $u(x) = 1 - e(-\beta x)$	1.37	0.02	—	17.29	102.41	20014	60.4

Note. Each entry is the median parameter estimate or median sub-index of fit. SUM refers to sum of squared deviations between judged strength of preference and predictions of model. $-\log(\Pi P)$ refers to negative logarithm of the product of probabilities of choices given the model. In S-CPT(4), β is fixed to 1.

Fit of Models to Individual Data

To fit the data for individuals according to the S-CPT models, judgments of strength of preference and choices were fit by the following model:

$$D(R, L) = a[\mathrm{U}(R) - U(L)] \tag{13}$$

$$P(R, L) = F[b(U(R) - U(L))] \tag{14}$$

$$U(R) = \Sigma u(x_i)[W(P_i) - W(P_{i-1})] \tag{15}$$

$$W(P) = \frac{cP^{\gamma}}{[cP^{\gamma} + (1 - P)^{\gamma}]} \tag{16}$$

$$u(x) = x^{\beta} \tag{17}$$

where $D(R, L)$ is the predicted judgment of the amount to be paid to receive gamble R over gamble L (including the sign indicating the direction of choice); $P(R, L)$ is the predicted probability of choosing R over L; $U(R)$ and $U(L)$ are the utilities of the gambles; a and b are constants; the outcomes have been ordered from highest to lowest, $x_1 > x_2 > \cdots x_n$; the probabilities, P_i, are decumulative probability that the outcome is greater than or equal to x_i; $u(x)$ is the utility of outcome, x; $W(P)$ is a cumulative weighting function, characterized by two constants, c and γ, F is the logistic function, $F[b(U(R) - U(L))] = 1/[1 + \exp(-b(U(R) - U(L))]$, where b is the spread parameter. There are five parameters to estimate: a, b, c, γ, and β.

To fit the P-CPT models, Eq. (16) was replaced with the following:

$$W(P) = P^{\gamma} \tag{18}$$

Two variations of P-CPT were fit, one using the power function of Equation 17 and the other using an exponential function for the utility function, as follows:

$$u(x) = c(1 - \exp(-\beta x)) \tag{19}$$

where c and β are constants. In this case, the constant c will be absorbed in a and b, so it can be set to 1 without loss of generality.

Models were fit to the 108 judgments (and choices) of nondominated gambles to minimize the following compromise of two criteria:

$$B = h \sum_{i=1}^{108} (D_i - \hat{D}_i)^2 - (1 - h) \log[\prod_{i=1}^{108} P(C_i)] \tag{20}$$

where D_i and $\hat{D}_i$ the observed and predicted judgment for each choice, which can be either positive or negative, depending on the direction of choice; C_i is the observed choice (gamble on the right or left), $P(C_i)$ is the probability of that choice given the model [i.e., $P(R, L)$ or $1 - P(R, L)$ from Equation 14]. The term, $\mathrm{SUM} = \Sigma_{i=1}^{108} (D_i - \hat{D}_i)^2$, is the familiar sum of squared deviations between observed and predicted judgments, and the term, $-\log[\Pi_{i=1}^{108} P(C_i)]$, is the (negative log) likelihood of the observed choices given the model; h is the weight given the two criteria. This compromise loss function requires the model to account for both the strength of preference judgments and choices using the same parameters. S-CPT and P-CPT models were fit using special programs, SUBFIT and LUCEFIT, which utilized Chandler's (1969) subroutine, STEPIT, to accomplish the minimizations. Similar results were obtained using BLACK-BOX instead of STEPIT, and similar results were also obtained when the value of h was set to different values between 0 and 1. The programs appeared to converge faster and perform better with h set to intermediate values than when h was set to zero or one. The values reported are for $h = .01$.

Median values of estimated parameters and the two sub-indices of fit are listed in Table 6. The best-fitting CPT model is S-CPT, with a power function for the utility function. Of 100 judges, 79 had estimated values of $\gamma > 1$, opposite the prediction of the inverse-S function, which would assume $\gamma < 1$. Of the 100 judges, 85 had estimates of $c < 1$, indicating greater weighting in 50-50 gambles for the lower outcome. The median value of the sum of squared deviations, SUM = 17,495, corresponds to a root mean squared deviation of \$12.73. The median negative log likelihood is 52.0, which is better than the value of 74.9 for the null model that sets all choice probabilities to 1/2. Of 100 judges, 60 had values of $\beta < 1$; fixing $\beta = 1$ made both indices of fit slightly worse, but this simplified S-CPT fit as well or better than both P-CPT models.

The simplified S-CPT model (with β fixed to 1) uses four parameters, the same as the P-CPT models. The P-CPT model with $u(x) = x^\beta$ fit slightly better than the simplified S-CPT model on the sum of squared deviations, but it was slightly worse predicting choices. Replacing the power function with the exponential utility function made the fit of the P-CPT model worse. This model fit worse than the simplified S-CPT model for 53 judges on both subindices, compared with only 19 who fit the P-CPT model with exponential utility function better on both subindices, and the other judges split on the two subindices. Both versions of P-CPT yield median estimates of $\gamma > 1$.

We also used the same procedures to fit a simplified version of the Birnbaum and Stegner (1979) model, using Equations 11 and 12 instead of 15 and 16 above, and fitting only one configural parameter in Equation 12, using the following simplification: $w(j, n) = \omega/(n + 1)$, where ω is the single configural parameter, and n is the number of outcomes in the gamble. This model was fit using $S(p) = p^\gamma$, and with $u(x) = x^\beta$, and it also used Equations 13, 14, and Equation 20. Like the S-CPT model, this model also has 5 parameters. The median estimates are as follows: $\gamma = 1.21$, $\beta = .61$, $\omega = -.45$, $a = .53$, and $b = 2.41$. The median sum of squared deviations was 16,446, slightly better than the best-fitting CPT model, and 53.6 for the negative log likelihood, slightly worse. When β was fixed to 1, the median subindices of fit were 18,864 and 57.0, comparable to the fits of the corresponding 4-parameter, CPT models. With β fixed to 1, the median estimates are $\gamma = 1.28$, $\omega = -.69$, $a = .075$, and $b = .50$. This model could be improved by using different values of $w(j, n)$, but that would complicate the comparison of fit by allowing additional parameters to one model but not the others.

DISCUSSION

The results of this experiment show systematic violations of branch independence and distribution independence. Violations of branch independence and distribution independence are inconsistent with EU, SEU, and SWU theories.

Systematic violations of branch independence and distribution independence are also inconsistent with the theory that judges consistently edit and eliminate common components prior to choosing between gambles (Kahneman & Tversky, 1979). While some judges might use such a strategy on some occasions, the theory that they do so consistently can be rejected.

The violations of branch independence and distribution independence observed are opposite the predictions of the inverse-S weighting function of cumulative prospect theory. The violations of branch independence show the same pattern as those obtained in previous studies of judgments and choices of gambles consisting of three and four equally likely outcomes (Birnbaum & McIntosh, 1996, submitted; Birnbaum & Beeghley, 1997; Birnbaum & Veira, in press; Weber & Kirsner, 1997). The present experiment used gambles with specified and varied levels of probability, so the present data indicate that the use of equally likely outcomes was not crucial to the previously observed pattern of results. All of these studies found evidence that weights have the opposite pattern from that of inverse-S cumulative prospect theory.

Wakker, Erev, and Weber (1994) failed to find systematic evidence of violations of branch independence predicted by rank-dependent models. Their study did not use "filler" gambles presented to judges on the notion that such trials would prevent the judge from learning that on every trial there is a common component that could be edited out.

However, Birnbaum and McIntosh (submitted) tested the idea that judges might learn to edit and cancel common outcomes if all of the experimental trials would permit such a cancellation. Birnbaum and McIntosh (submitted) removed the "filler" trials that had been used in their earlier studies, and found that violations of branch independence were reduced slightly, but they were similar to their previously obtained results. Thus, the lack of "fillers" does not completely explain the null finding of Wakker *et al.* (1994). Weber and Kirsner (1997) used a variant of the design of Wakker *et al.* and found small, but significant violations in the same direction as observed by Birnbaum and McIntosh (1996). The Wakker, *et al.* study did not use the systematic variation of the terms in Eqs. (3b) and (4b), as in the designs of Birnbaum and McIntosh (1996; submitted), or as in Table 1; consequently, the earlier design of Wakker *et al.* (1994) may have "missed" finding outcomes for which the weights would straddle the ratio of differences in utility for most judges.

The violations of distribution independence observed here are a new result, and these violations are inconsistent with the model of Birnbaum and McIntosh

(1996), which implies that there should be no violations of this property. Violations of distribution independence are also inconsistent with the extension of original prospect theory (Kahneman & Tversky, 1979) to gambles with four outcomes.

Both patterns of violations are consistent with cumulative prospect theory, assuming a cumulative weighting function that is S-shaped, rather than inverse-S shaped. Other weighting functions satisfying the property of Expression (3b) can also be retained, as long as they can also account for the violations of distribution independence. The power function (Expression (5c)) can have this property when $\gamma > 1$, but P-CPT did not fit the data of the majority of individual judges as well as S-CPT. P-CPT requires weights of different rank positions to be a monotonic function of rank (for equally likely outcomes, as in Subdesign 1). The better fit of the S-CPT model may be due to its flexibility to allow weights of medium outcomes to be greater than weights of extreme outcomes. Birnbaum and Beeghley (1997) and Birnbaum and Veira (in press) concluded that medium level outcomes have greater weights than extremes in the seller's point of view, if the utility function is assumed to be linear. Nevertheless, it is possible that with some other utility function besides the power function or exponential, P-CPT might achieve a better fit, so this theory cannot be rejected based on the present data alone.

The present data contradict the inverse-S weighting function observed in experiments by Tversky and Kahneman (1992) and Wu and Gonzalez (1996). There are two ways to explain this apparent contradiction: (1) one could assume that CPT is correct and conclude that the contradiction between the S and inverse-S weighting functions is evidence that the weighting function induced by the context in this experiment is different from that induced in certain other experiments, or (2) one could take the contradiction as evidence against CPT theory. As noted by Birnbaum (1997) and Birnbaum and McIntosh (1996; submitted), configural weight theories can explain both sets of data with the same parameters, whereas CPT requires that the weighting function change to account for the data of both Birnbaum and McIntosh (1996) and Wu and Gonzalez (1996). We favor the second interpretation, that CPT is flawed, based on simplicity (one weighting function instead of two) and converging evidence from recent experiments testing stochastic dominance and cumulative independence (Birnbaum & Navarrete, submitted), described below.

Although the Birnbaum and McIntosh (1996) model is refuted by systematic violations of distribution independence, the "revised" configural weight model of Birnbaum and Stegner (1979) remains consistent with the present results, as does the minimum loss function theory of Birnbaum, *et al.* (1992; see also Birnbaum & McIntosh, 1996). The simplified configural model fit here, like the P-CPT model, assumes that equally likely outcomes will have their weights monotonically related to ranks. In order to account for the effects of viewpoint in judgment studies, however, it may be necessary to allow the configural weight parameters to follow a more complex form than that imposed here on Eq. (12). Properties that can distinguish Cumulative Prospect theories (as a class) from the Configural weight models and from minimum loss theory have been discussed by Birnbaum (1997). Three of these properties are comonotonic branch independence, stochastic dominance, and cumulative independence.

This experiment did not test comonotonic independence (a requirement of the generic rank-dependent, configural weight model used here). All versions of CPT and the configural weight models of both Birnbaum and McIntosh (1996) and Birnbaum and Stegner (1979) imply comontonic independence when the number of outcomes is held constant. Minimum loss theory, however, can lead to a configural weight model that can violate comonotonic independence (Birnbaum & McIntosh, 1996, Appendix A). In the loss theory model, weights not only vary as a function of the rank of the outcomes, but they also depend on the spacing of the outcomes, which allows minimum loss theory to violate global comonotonic independence, though it satisfies comonotonic independence in local regions.

Birnbaum (1997) noted that configural weighting models can violate stochastic dominance, whereas all versions of CPT must satisfy this property. Stochastic dominance requires that if $P(x > t|A) \geq P(x > t|B)$ for all t, then gamble B should be not be preferred to gamble A. Birnbaum and Navarette (submitted) found that stochastic dominance was systematically violated in cases predicted by the Birnbaum and McIntosh (1996) model and parameters. For example, 73 out of 100 individuals tested chose B = ($12, .10; $90, .05; $96, .85) over A = ($12, .05; $14, .05; $96, .90), even though A stochastically dominates B. The mean judgment of the amount offerred to receive B instead of A, with positive numbers reflecting violations of dominance and negative numbers representing satisfaction of dominance, was $9.40! Similar results were observed for other choices, constructed from the same recipe.

There are two cumulative independence conditions that put the apparent contradiction in weighting functions between Wu and Gonzalez (1996) and Birnbaum and McIntosh (1996) into a single experiment. As shown in Birnbaum (1997), cumulative independence is a combination of comonotonic branch independence, monotonicity, transitivity, and coalescing equivalences. Coalescing is the assumption that a three-outcome gamble in which two outcomes are equal is equivalent to the two-outcome gamble in which the probabilities of the two equal outcomes are added. Cumulative prospect theory implies coalescing, whereas the configural weight theories do not. Birnbaum and Navarrete (submitted) found systematic violations of cumulative independence, violating any version of the CPT models, but predicted by the Birnbaum and McIntosh (1996) model.

Violations of stochastic dominance and cumulative independence suggest that the apparent contradiction between the S and inverse-S weighting functions in different experiments is due to a flaw in CPT, rather than a changing weighting function between studies, because CPT with any weighting function must satisfy stochastic dominance and cumulative independence.

The Birnbaum and Stegner (1979) model violates asymptotic independence, unlike the Birnbaum and McIntosh (1996) model. In a two-outcome gamble, as the probability of the higher outcome approaches one, the value of the gamble may not asymptote at the same level, independent of the value of the lowest outcome. Such violations of asymptotic independence were observed in morality judgment by Birnbaum (1973) and Riskey and Birnbaum (1974): Given a person has done one very bad deed, the person will be judged "immoral" no matter how many good deeds the person has also done. For decision making, the most direct analog is to insurance. Given a chance of a bad outcome, there may be no value of the probability that makes the utility of the gamble asymptote to the value it would if the worst outcome could be changed into a good outcome. From the viewpoint of the Birnbaum and Stegner (1979) model, the purpose of insurance is to change the value of the worst outcome.

In conclusion, choices between three- and four-outcome gambles are inconsistent with EU, SEU, SWU, and the theory that judges consistently edit and eliminate common components. Violations of branch independence and distribution independence are opposite the predictions of the inverse-S weighting function of the Tversky and Kahneman (1992) cumulative prospect model. Violations of distribution independence are not consistent with either the model of Birnbaum and McIntosh (1996) or an extension of original prospect theory to four-outcome gambles. Data remain consistent with cumulative prospect theory if the weighting function satisfies Expression (3b), rather than Expression (4b). Data were better fit by the S-shaped cumulative weighting function than by a power weighting function, but the power weighting function is not qualitatively eliminated. Data also remain consistent with two other theories that can violate distribution independence and branch independence: Birnbaum and Stegner's (1979) revised configural weighting model and minimum loss theory.

REFERENCES

Birnbaum, M. H. (1973). Morality judgment: Test of an averaging model with differential weights. *Journal of Experimental Psychology,* **99,** 395–399.

Birnbaum, M. H. (1974). The nonadditivity of personality impressions. *Journal of Experimental Psychology,* **102,** 543–561.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Eds.), *Social attitudes and psychophysical measurement* (pp. 401–485). Hillsdale, NJ: Erlbaum.

Birnbaum, M. H. (1992). Issues in utility measurement. *Organizational Behavior and Human Decision Processes,* **52,** 319–330.

Birnbaum, M. H. (1997). Violations of monotonicity in judgment and decision making. In A. A. J. Marley (Ed.), *Choice, decision and measurement: Essays in honor of R. Duncan Luce* (pp. 73–100). Mahwah, NJ: Erlbaum.

Birnbaum, M. H., & Beeghley, D. (1997). Violations of branch independence in judgments of the value of gambles. *Psychological Science,* **8,** 87–94.

Birnbaum, M. H., Coffey, G., Mellers, B. A., & Weiss, R. (1992). Utility measurement: Configural-weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 331–346.

Birnbaum, M. H., & Jou, J. W. (1990). A theory of comparative response times and "difference" judgments. *Cognitive Psychology,* **22,** 184–210.

Birnbaum, M. H., & Navarrete, J. *Testing rank- and sign-dependent utility theories: Violations of stochastic dominance and cumulative independence.* Manuscript submitted for publication.

Birnbaum, M. H., & McIntosh, W. R. (1996). Violations of branch independence in choices between gambles. *Organizational Behavior and Human Decision Processes,* **67,** 91–110.

Birnbaum, M. H., & McIntosh, W. R. *Violations of branch independence without fillers: Evidence against cancellation theory.* Manuscript submitted for publication.

Birnbaum, M. H., & Sotoodeh, Y. (1991). Measurement of stress: Scaling the magnitudes of life changes. *Psychological Science,* **2,** 236–243.

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology,* **37,** 48–74.

Birnbaum, M. H., & Stegner, S. E. (1981). Measuring the importance of cues in judgment for individuals: Subjective theories of IQ as a function of heredity and environment. *Journal of Experimental Social Psychology,* **17,** 159–182.

Birnbaum, M. H., & Sutton, S. E. (1992). Scale convergence and utility measurement. *Organizational Behavior and Human Decision Processes,* **52,** 183–215.

Birnbaum, M. H., Thompson, L. A., & Bean, D. J. (1997). Testing interval independence versus configural weighting using judgments of strength of preference. *Journal of Experimental Psychology: Human Perception and Performance,* **23,** 939–947.

Birnbaum, M. H., & Veira, R. (in press). Configural weighting in two- and four-outcome gambles. *Journal of Experimental Psychology: Human Perception and Performance.*

Birnbaum, M. H., & Zimmermann, J. M. *Configural weighting model of buying and selling prices predicts violations of joint independence in judgments of investments.* Manuscript submitted for publication.

Champagne, M., & Stevenson, M. K. (1994). Contrasting models of appraisal judgments for positive and negative purposes using policy modeling. *Organizational Behavior and Human Decision Processes,* **59,** 93–123.

Chandler, J. P. (1969). STEPIT: Finds local minima of a smooth function of several parameters (CPA 312). *Behavioral Science,* **14,** 81–82.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin,* **51,** 380–417.

Edwards, W. (1992). Towards the demise of economic man and woman: Bottom lines from Santa Cruz. In W. Edwards (Eds.), *Utility theories: Measurements and applications* (pp. 254–267). Boston, MA: Kluwer.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica,* **47,** 263–291.

Karmarkar, U. S. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance,* **21,** 61–72.

Lopes, L. (1990). Re-modeling risk aversion: A comparison of Bernoullian and rank dependent value approaches. In G. M. von Furstenberg (Eds.), *Acting under uncertainty* (pp. 267–299). Boston, MA: Kluwer.

Luce, R. D. (1992). Where does subjective expected utility fail descriptively? *Journal of Risk and Uncertainty,* **5,** 5–27.

Luce, R. D. (1996). When four distinct ways to measure utility are the same. *Journal of Mathematical Psychology,* **40,** 297–317.

Luce, R. D., & Fishburn, P. C. (1991). Rank- and sign-dependent linear utility models for finite first order gambles. *Journal of Risk and Uncertainty,* **4,** 29–59.

Luce, R. D., & Fishburn, P. C. (1995). A note on deriving rank-dependent utility using additive joint receipts. *Journal of Risk and Uncertainty,* **11,** 5–16.

Miyamoto, J. M. (1989). Generic utility theory: measurement foundations and applications in multiattribute utility theory. *Journal of Mathematical Psychology,* **32,** 357–404.

Prelec, D. (1995). *The probability weighting function.* Working paper, Sloan School of Management, MIT, E56-320, 38 Memorial Drive, Cambridge, MA 02139.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization,* **3,** 324–345.

Riskey, D. R., & Birnbaum, M. H. (1974). Compensatory effects in moral judgment: Two rights don't make up for a wrong. *Journal of Experimental Psychology,* **103,** 171–173.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Starmer, C., & Sugden, R. (1989). Violations of the independence axiom in common ratio problems: An experimental test of some competing hypotheses. *Annals of Operations Research,* **19,** 79–101.

Stevenson, M. K., Busemeyer, J. R., & Naylor, J. C. (1991). Judgment and decision-making theory. In M. Dunnette & L. M. Hough (Eds.), *New handbook of industrial-organizational psychology* (pp. 283–374). Palo Alto, CA: Consulting Psychologist Press.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review,* **79,** 281–299.

Tversky, A., & Fox, C. R. (1995). Weighing Risk and Uncertainty. *Psychological Review,* **102**(2), 269–283.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business,* **59,** S251–S278.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty,* **5,** 297–323.

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica,* **63,** 1255–1280.

Viscusi, K. W. (1989). Prospective reference theory: Toward an explanation of the paradoxes. *Journal of risk and uncertainty,* **2,** 235–264.

Wakker, P., Erev, I., & Weber, E. U. (1994). Comonotonic independence: The critical test between classical and rank-dependent utility theories. *Journal of Risk and Uncertainty,* **9,** 195–230.

Wakker, P., & Tversky, A. (1993). An axiomatization of cumulative prospect theory. *Journal of Risk and Uncertainty,* **7,** 147–176.

Weber, E. U., & Kirsner, B. (1997). Reasons for rank-dependent utility evaluation. *Journal of Risk and Uncertainty,* **14,** 41–62.

Weber, E. U. (1994). From subjective probabilities to decision weights: The effects of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin,* **114,** 228–242.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science,* **42,** 1676–1690.

Received: December 30, 1996