Judgment and Decision Making, Vol. 3, No. 7, October 2008, pp. 528–546

Tests of Cumulative Prospect Theory with graphical displays of probability

Michael H. Birnbaum*, Kathleen Johnson, and Jay-Lee Longbottom
California State University, Fullerton

Abstract

Recent research reported evidence that contradicts cumulative prospect theory and the priority heuristic. The same body of research also violates two editing principles of original prospect theory: cancellation (the principle that people delete any attribute that is the same in both alternatives before deciding between them) and combination (the principle that people combine branches leading to the same consequence by adding their probabilities). This study was designed to replicate previous results and to test whether the violations of cumulative prospect theory might be eliminated or reduced by using formats for presentation of risky gambles in which cancellation and combination could be facilitated visually. Contrary to the idea that decision behavior contradicting cumulative prospect theory and the priority heuristic would be altered by use of these formats, however, data with two new graphical formats as well as fresh replication data continued to show the patterns of evidence that violate cumulative prospect theory, the priority heuristic, and the editing principles of combination and cancellation. Systematic violations of restricted branch independence also contradicted predictions of "stripped" prospect theory (subjectively weighted additive utility without the editing rules).

Keywords: Cumulative Prospect Theory, TAX model, priority heuristic, graphical format, branch independence, cancellation, combination, choice, uncertainty.

1 Introduction

Probably the most popular descriptive model of risky decision making in the last fifteen years has been cumulative prospect theory (CPT) (Camerer, 1998; Starmer, 2000; Tversky & Fox, 1995; Tversky & Kahneman, 1992; Wu, Zhang, & Gonzalez, 2004). CPT provides a way to describe the Allais paradoxes, which are findings that lead expected utility (EU) theory into self-contradiction.

However, recent findings have been reported that contradict CPT (Birnbaum, 1999b, 2001, 2004a, 2004b, 2005, 2006; 2008b; Birnbaum & Navarrete, 1998; Weber, 2007). These findings have been described as "new paradoxes" because CPT is forced into self-contradiction when it attempts to use a probability weighting function and value (utility) function to analyze these results (Birnbaum, 2008b). These violations of CPT were predicted by Birnbaum's transfer of attention exchange (TAX) model, which was used to design the new tests (Birnbaum, 1997, 1999a, 1999b; 2008b; Birnbaum & Chavez, 1997; Birnbaum & Navarrete, 1998).

Many of the same tests that violate CPT also contradict the priority heuristic (Birnbaum, 2008a; b, c; Brandstätter, Gigerenzer, & Hertwig, 2006).

1.1 Five violations of CPT and PH

Five properties implied by CPT and tested in these studies are listed in Table 1, along with examples of choices that have shown significant and systematic violations in previous research. These phenomena also contradict the predictions of the priority heuristic. If these phenomena are "real," it means that we cannot regard CPT or PH as accurate descriptive models of risky decision making. Appendices A, B, and C describe CPT, PH, and TAX models and their predictions for the five tests in Table 1.

However, there may be a way to salvage CPT or PH (at least in a limited way) if we could find some set of procedures that lead to data that are compatible with those theories.

1.2 Different procedures can yield different preference orders

Different procedures for eliciting preferences in risky decision-making can result in different preference orderings. With one procedure, we find that A is preferred to B, and with another procedure, we find that B is preferred to A. For example, people may set a higher price to sell A than they do to sell B but when given a choice, they may prefer B over A (Lindman, 1971, Lichtenstein & Slovic,

*We thank Jeffrey P. Bahra for assistance in testing the participants.
Address: Prof. Michael H. Birnbaum, Department of Psychology, CSUF H-830M, P.O. Box 6846, Fullerton, CA 92834–6846. Email: mbirnbaum@fullerton.edu.

Table 1: Properties tested in the experiments, including five "new paradoxes" that violate CPT. (S $\prec$ R denotes S is preferred to R.)

Property	Expression	Example Violation of CPT
Stochastic Dominance[a]	$G^+ = (x, p; y^+, q'; y, r - q') \succ$ $G^- = (x, p - q; x^-, q; y, r)$	$G^+ =$ (\$96, 0.90; \$14, 0.05; \$12, 0.05) $\prec$ $G^- =$ (\$96, 0.85; \$90, 0.05; \$12, 0.10)
Coalescing	$G^+ \succ G^- \Leftrightarrow GS^+ \succ GS^-$	$G^+ \prec G^-$ and $GS^+ =$ (\$96, 0.85; \$96, 0.05; \$14, 0.05; \$12, 0.05) $\succ$ $GS^- =$ (\$96, 0.85; \$90, 0.05; \$12, 0.05; \$12, 0.05)
Lower Cumulative Independence[b]	$S = (x, p; y, q; z, r) \succ R = (x', p; y', q; z, r)$ $\Rightarrow S'' = (x, p + q; y', r) \succ R'' = (x', p; y', q + r)$	$S =$ (\$44, 0.1; \$40, 0.1; \$2, 0.8) $\succ$ $R =$ (\$98, 0.1; \$10, 0.1; \$2, 0.8) and $S'' =$ (\$44, 0.2; \$10, 0.8) $\prec$ $R'' =$ (\$98, 0.1; \$10, 0.9)
Upper Cumulative Independence[b]	$S' = (z', r; x, p; y, q) \prec R' = (z', r; x', p; y', q)$ $\Rightarrow S''' = (x', r; y, p + q) \prec R''' = (x', p + r; y', q)$	$S' =$ (\$110, 0.8; \$44, 0.1; \$40, 0.1) $\prec$ $R' =$ (\$110, 0.8; \$98, 0.1; \$10, 0.1) and $S''' =$ (\$98, 0.8; \$40, 0.2) $\succ$ $R''' =$ (\$98, 0.9; \$10, 0.1)
Restricted Branch Independence[b, c]	$S = (x, p; y, p; z, r) \succ R = (x', p; y', p; z, r) \Leftrightarrow$ $S' = (z', r; x, p; y, p) \succ R' = (z', r; x', p; y', p)$	$S =$ (\$44, 0.1; \$40, 0.1; \$2, 0.8) $\succ$ $R =$ (\$98, 0.1; \$10, 0.1; \$2, 0.8) $S' =$ (\$110, 0.8; \$44, 0.1; \$40, 0.1) $\prec$ $R' =$ (\$110, 0.8; \$98, 0.1; \$10, 0.1)

Notes: [a] $p + r = 1; 0 < q < p; 0 < q' < r; 0 < y < y^+ < x^- < x$.
[b] $p + q + r = 1; 0 < z < y' < y < x < x' < z'$.
[c] $2p + r = 1$.

1971; Goldstein & Einhorn, 1987; Birnbaum & Sutton, 1992; Slovic & Lichtenstein, 1983; Tversky, Slovic, & Kahneman, 1990). Such results show that ordering of gambles from judgments and the ordering inferred from choices are different.

In fact, two types of judgments need not produce the same preference order. For example, buying and selling prices of used cars, risky gambles and stock investments are not monotonically related to each other (Birnbaum & Stegner, 1979; Birnbaum & Sutton, 1992; Birnbaum & Zimmermann, 1998): the buying price for *A* can be greater than the buying price for *B* and yet people ask more to sell *B* than to sell A. Attractiveness ratings and judged buying prices are also not monotonically related to each other (Tversky, Sattath & Slovic, 1988). Furthermore, the ordering of attractiveness ratings can be manipulated by changing the context of other gambles that are presented in the same study (Mellers, Ordóñez, & Birnbaum, 1992).

People may also reverse choices when consequences in gambles are described ("framed") differently. For example, people have been asked if they would prefer to have \$40 for sure or a 50–50 gamble to win either \$100 or \$0. Most people chose the \$40 for sure. However, when the same people are given \$100 contingent on accepting one of two losing gambles, most preferred a 50–50 gamble to lose the \$100 or lose nothing, rather than accept a sure loss of \$60. But both situations lead to the same final consequences, because losing \$60 for sure from a \$100 endowment is the same as gaining \$40 for sure (Kahneman & Tversky, 1979; Birnbaum, 2001).

In addition, preference orders may differ when probability or uncertainty is learned from experience as opposed to described verbally (Bleaney & Humphrey, 2006; Hertwig, Barron, Weber, & Erev, 2004; but see also Fox & Hadar, 2006).

1.3 Form, format, and framing

The fact that results can be altered by changing procedures has led some to question whether results in the field of decision making will generalize from one situation to another. To describe these manipulations of procedure, the terms "framing" and "representation" have been used to describe many different variables (Harless, 1992; Keller, 1985; Tversky & Kahneman, 1986). Levin, Schneider, & Gaeth (1998) presented a taxonomy of different types of "framing" effects. Birnbaum (2004b) suggested that we use different terms to distinguish three variables of procedure: form, format, and framing.

The *form* of a gamble refers to the manner in which branches are split or coalesced. For example, the gamble A = ($100, .02; $0, 0.98) is a two-branch gamble with a probability of 0.02 to win $100 and otherwise receive $0; the gamble A' = ($100, 0.01; $100, 0.01; $0, 0.98) is a three-branch split form of A in which there are two branches with probability of 0.01 to win $100 and otherwise $0. Clearly, there are many other split forms of A.

Kahneman and Tversky (1979) theorized that form should not affect the evaluation of gambles, so they proposed that people combine branches leading to the same consequence and convert A' to A before evaluating it or comparing it with other gambles (see Kahneman, 2003). Cumulative prospect theory (Tversky & Kahneman, 1992) automatically satisfies this property (that form has no effect).

Format refers to how the probability mechanism, consequences, and probabilities are represented to participants, including the manner in which a choice is displayed. *Framing* refers to how consequences and events are described. Whereas Kahneman and Tversky (1979) assumed that form should not affect choices, an implication of cumulative prospect theory, they opined that format and framing would have large effects (Tversky & Kahneman, 1986).

Because of numerous findings that judgments and decisions depend on these various aspects of procedure, whenever a new phenomenon is unearthed, investigators are naturally suspicious that the new results might be altered or even reversed by a change of procedure. For example, Harless (1992) concluded that effects described as "regret" effects did not occur with certain formats for presentation of gambles (he used the term "representation" to refer to format). However, the study of Harless confounded form and format; his effects were later attributed to event-splitting effects (form) rather than to regret effects or representation (format) effects (Birnbaum, 2006; Humphrey, 1995; Starmer & Sugden, 1993).

1.4 Purpose of the Present Studies

The studies in this article were conducted to resolve two collaborative disputes, concerning whether formats could be devised that would produce data that satisfy principles of prospect theories. New formats for the presentation of gambles were devised with the intention to reverse results that had previously violated stochastic dominance and coalescing and which produced violations of restricted branch independence that are opposite those predicted by the inverse-S shaped weighting function of CPT needed to account for the Allais paradoxes. Coalescing and stochastic dominance must be satisfied according to either CPT or to original prospect theory with its editing rules.

Because CPT had been widely regarded as an accurate descriptive model, it is important to determine if the experiments that refute it might yield different conclusions if procedures were changed. The second and third coauthors of this paper devised graphical formats that they were confident would alter the results of Birnbaum (1999b; 2004a; 2004b). While it is agreed that a picture can be worth a thousand words, it is also clear that graphs can be manipulated to create very different impressions of a given set of data (Huff, 1954). What is not yet known is how to present information in the "best" way possible (Lipkus, 2007; Cutie, Weinstein, Emmons, & Colditz, 2008). It was conjectured that certain graphical formats would make it easy for people to perceive first order stochastic dominance, a property that most decision makers regard as rational, and which is implied by CPT.

1.5 Violations of Stochastic Dominance

Birnbaum and Navarrete (1998) and Birnbaum (1999b) reported that about 70% of undergraduates violate first order stochastic dominance when asked to choose between gambles such as the following. Suppose there are two urns, each of which contains 100 tickets. You can choose the urn from which a ticket will be drawn at random and the amount printed on that ticket will determine your prize. Which urn do you choose?

I:	90 tickets to win $96
	5 tickets to win $14
	5 tickets to win $12
J:	85 tickets to win $96
	5 tickets to win $90
	10 tickets to win $12

Most undergraduates tested with a format like this violate first order stochastic dominance by selecting gamble J over gamble I, even though I dominates J. Note that the probability to win $96 or more is higher in I than J; the probability to win $14 or more is higher in I than J, and the probabilities to win $90 or more and $12 or more are the same in both gambles. According to CPT, people should choose I over J. This prediction holds under CPT with any probability weighting function and with any strictly increasing value function for the consequences.

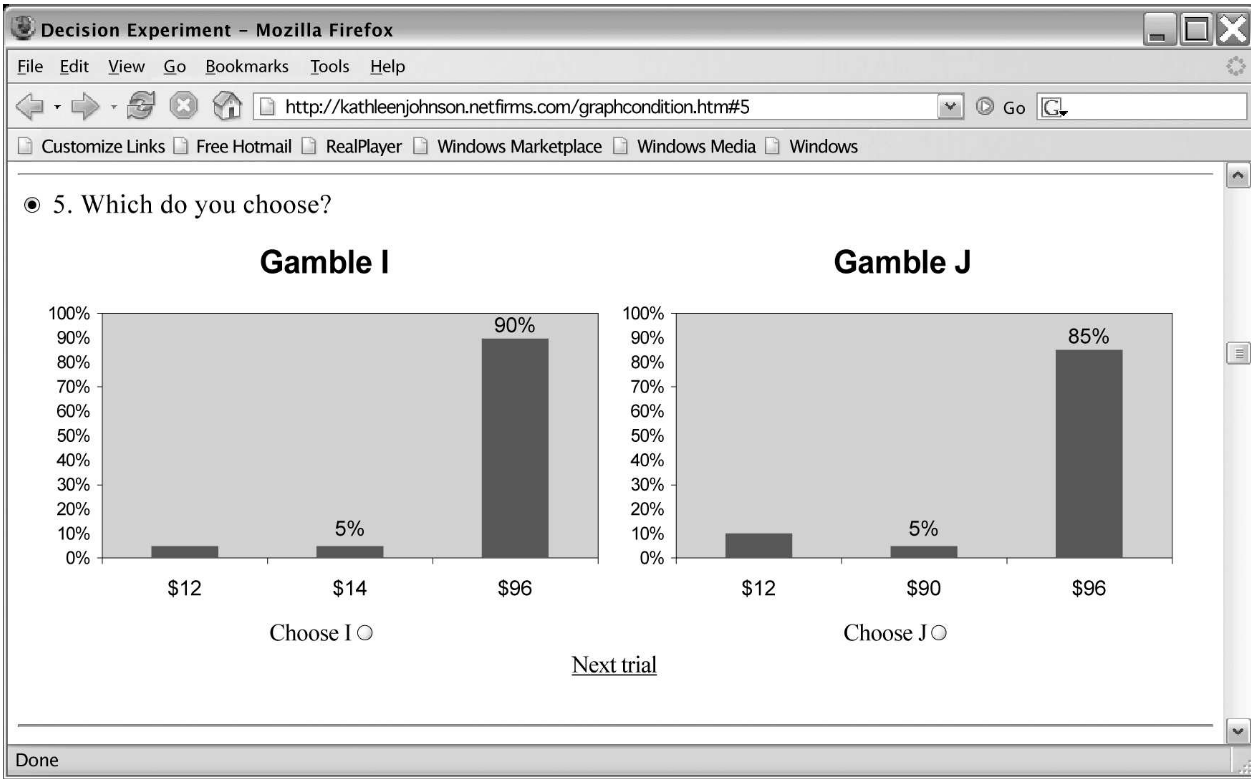

Figure 1: Example of a choice in histogram format. It was thought that people could see that the probability to win the highest prize is higher in Gamble *I* than *J* and the probability to receive the worst consequence is lower in *I* than *J*. Gamble *I* dominates *J*, but the majority of participants chose *J* over *I*.

It was conjectured that if this choice were illustrated by means of histograms representing the gambles, most people would be able to "see" dominance visually. For example, in Figure 1, a person should be able to see that 10% to win $12 in the gamble on the right (Gamble *J*) has been replaced by 5% to win $12 and 5% to win $14 on the left (Gamble *I*); furthermore, the 90% to win $96 in *I* has been replaced by 85% to win $96 and 5% to win $90. Birnbaum (2004b) used pie charts to represent probabilities in gambles and found that the majority continued to violate stochastic dominance. However, it was conjectured that histograms (bar charts as in Figure 1) are more familiar displays (than pie charts), and that the representation of probability by height would reveal dominance immediately and visually. This hypothesis is tested in the first study using histograms to represent the gambles. It is also tested in Study 2, in which each equally likely consequence was presented in a list format that also contained a histogram-like arrangement in which the height of a list was proportional to relative frequency (Figure 2).

1.6 Allais Paradoxes

Another source of evidence against CPT is evidence found in a dissection of the Allais constant consequence paradoxes. Birnbaum (1999a; 2004a; 2008b) noted that the constant consequence paradox of Allais might be due to violations of restricted branch independence (according to CPT), to violations of coalescing (according to original prospect theory), or to both (according to the TAX model). By dissecting the Allais paradox experimentally (allowing separate tests of restricted branch independence and coalescing), it is possible to compare these theories.

It can be shown that expected utility theory implies that a person should satisfy the following independence property (see Appendix D):

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

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

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

where $x > y > z \geq 0$, $1 \geq p > q \geq 0$, and all probabilities are between 0 and 1, and $S \succ R$ denotes that *S* is preferred to *R*. Gamble *S* is called the "safe" alternative because it has a higher probability of winning a smaller prize whereas *R* is termed "risky" because it has a smaller probability to win a greater prize. Violations of these properties, $S \succ R \Leftrightarrow S'' \succ R''$, $S'' \succ R'' \Leftrightarrow S' \succ R'$, and $S \succ R \Leftrightarrow S' \succ R'$ are termed *type 1, type 2,* and *type 3 Allais paradoxes*, respectively. They constitute evidence against expected utility theory (Appendix D). Both TAX and CPT imply violations of these properties, but for different reasons.

For example, suppose there are two urns that contain 100 tickets each: a "safe" alternative, *S,* and a "risky" one, *R*. You can choose from which urn to draw a random ticket and the value printed on that ticket is your prize. Consider the following two choices:

Choice 1: S = (\$40, .2; \$2, 0.8) versus R = (\$98, 0.1; \$2, 0.9). This is a choice between 20 tickets to win \$40 (and 80 to receive \$2) and 10 tickets to win \$98 (and 90 to receive \$2).

Now consider Choice 2: S' = (\$98, 0.8; \$40, 0.2) versus R' = (\$98, 0.9; \$2, 0.1). This is a choice between an 80% chance to receive \$98 (otherwise receive \$40) and a 90% chance to win \$98 (otherwise receive \$2). According to expected utility theory, a person should either choose both "safe" gambles or choose both "risky" gambles, but should not switch.

Because many people choose R over S and S' over R', contrary to expected utility, the result is called "paradoxical." Birnbaum (2004a) noted that if people satisfied *transitivity* [$A \succ B$ and $B \succ C \Rightarrow A \succ B$], *coalescing* [$A = (x, p; y, 1 - p) \sim A' = (x, p - q; x, q; y, 1 - p) \sim A'' = (x, p; y, 1 - p - r; y, r)$], and *restricted branch independence* [$S = (x, p; y, q; z, 1 - p - q) \succ R = (x', p; y', q; z, 1 - p - q) \Leftrightarrow S' = (z', 1 - p - q; x, p; y, q) \succ R' = (z', 1 - p - q; x', p; y', q)$], they would not show such reversals (Allais paradoxes). Therefore, any transitive theory must violate either restricted branch independence or coalescing to account for these Allais paradoxes. Original prospect theory satisfies branch independence and violates coalescing (apart from its editing principles), whereas CPT satisfies coalescing and violates restricted branch independence (apart from editing). See Appendix E. The TAX model and PH violate both of these properties, and PH violates transitivity as well.

Birnbaum (2004a, 2007) conducted experimental dissections of Allais paradoxes and concluded that Allais paradoxes are due to violations of coalescing. Violations of restricted branch independence are also observed, but these are opposite the pattern required to explain the Allais paradoxes. Results of Birnbaum (2004a) were correctly predicted by TAX using prior parameters. As noted by Brandstätter et al. (2006), the PH does not account for those results.

1.7 List format might promote cancellation and combination

If people used a cancellation strategy in which components that are the same in both gambles are cancelled prior to making a decision (Kahneman & Tversky, 1979), then they would satisfy restricted branch independence. If people used the combination editing principle of original prospect theory, then they would satisfy coalescing. Both of these principles are contradicted by results of Birnbaum (2004a, 2004b, 2007).

However, it was hypothesized that if gambles were presented as lists of equally likely consequences, and if these were arranged with a vertical alignment of the consequences, it should be easy for people to see and to cancel common consequences. This list format also creates the appearance of histograms, but it is even more concrete in the sense that each equally likely outcome is also represented individually. If one wanted to, it would be easy to cross out tickets that are the same in both urns and to decide based on those tickets that remain.

If people cancelled common consequences in this way, they would satisfy stochastic dominance. For example, in Figure 2, a person should be able to simplify the choice by canceling the 17 tickets to win \$96, which are the same in both I and J; similarly, one could cancel the ticket to win \$12 common to both. That would leave \$96 and \$14 in choice *I,* compared with \$90 and \$12 in choice *J*, making it (in theory) easy to see that I dominates J in this display.

Similarly, if people cancelled common consequences, they would satisfy restricted branch independence, as should be apparent from Figure 3. If one were to cancel the common branch of 16 tickets to win \$98, then one should make the same choice as in other choices that differ only in the consequence on those sixteen common tickets. In other words, if the 16 tickets to win \$98 were changed to 16 tickets to win \$2, the choice should be the same because the common consequences would be cancelled in both cases.

2 Methods

In both studies, participants viewed the materials via computers connected to the WWW and clicked buttons to indicate the gamble in each choice that they would prefer to play. There were 20 choices in each study. In the histogram and text study, the choices were the same as those used in Birnbaum (1999b; 2004b); in the vertical list study, choices were the same as in Birnbaum (2004a). Gambles were described in terms of urns containing tickets that were equally likely but which had different prize values. The prize would be determined by a random draw from the selected urn.

2.1 Study 1: Histograms and text displays

Participants were randomly assigned to two conditions that used different formats to display gambles. Twelve of the choices provided two tests each of stochastic dom-

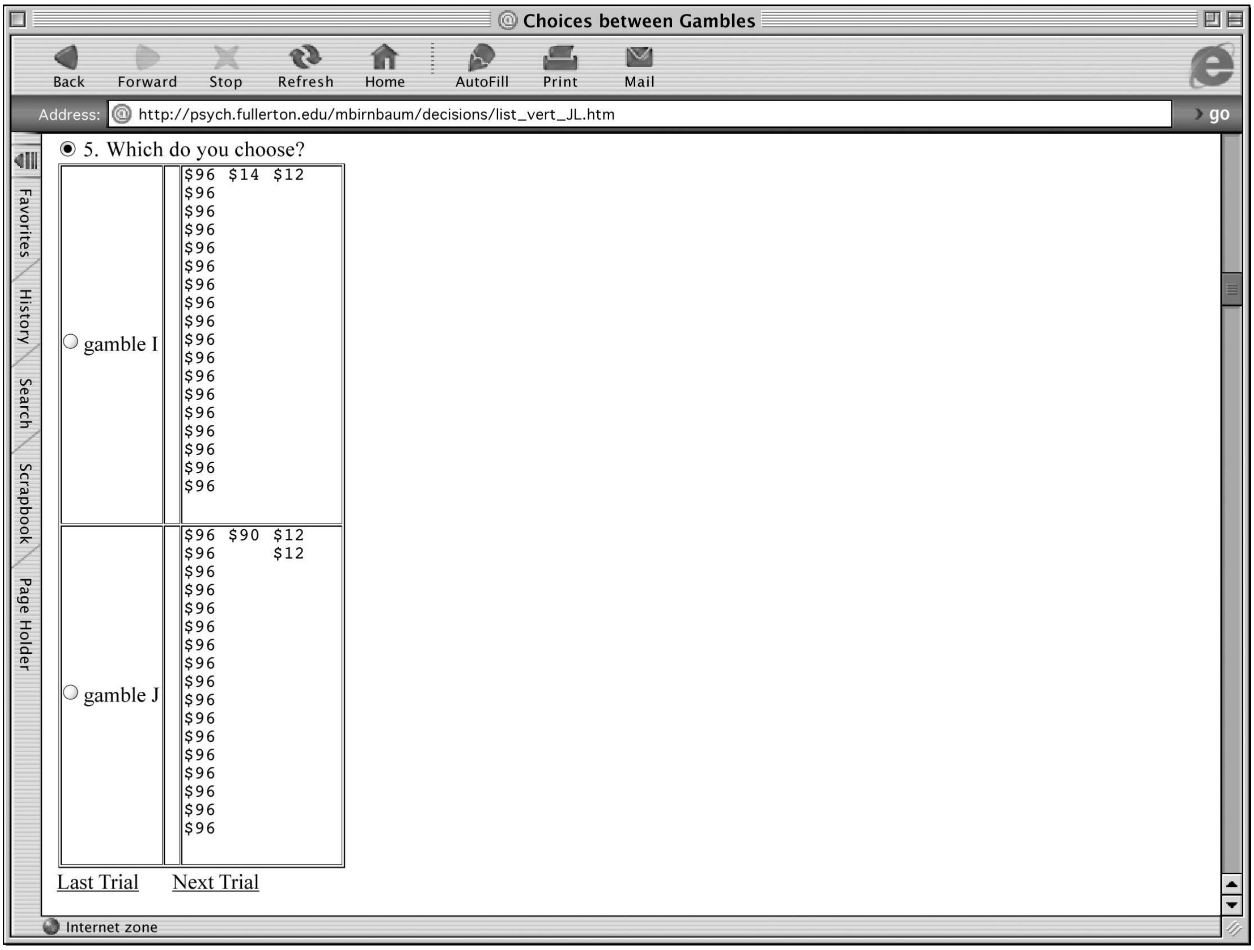

Figure 2: Example presentation of one trial in the vertical list format. By canceling equal tickets from both gambles, we are left with $96 and $14 in *I*, and with $90 and $12 in *J*, so this format should make it easy to see that *I* dominates *J*. Nevertheless, most chose *J*.

inance (choices 5 and 7), coalescing (choices 5 vs. 11 and 7 vs. 13), branch independence (6 vs. 10 and 17 vs. 12), lower cumulative independence (6 vs. 8 and 17 vs. 20), upper cumulative independence (10 vs. 9 and 12 vs. 14), and there were six other trials that served as warm-ups and fillers, the same as in Birnbaum (1999b).

In the *histogram* condition, choices were displayed as in Figure 1. Participants were instructed that the height of each bar in each histogram represents the probability of winning each amount of money given below that bar, the taller the bar the more likely to win that prize. Note that in addition to the graphic displays, percentages were displayed in the figures.

The *text* condition listed branches in ascending order of value of consequences, as in Birnbaum (1999b). Probability was described as the ratio of number of each type of ticket to the total number of tickets in the urn, which was always 100. Choices were displayed as in the following example:

5. Which do you choose?

I:	.05 probability to win $12
	.05 probability to win $14
	.90 probability to win $96
OR	
J:	.10 probability to win $12
	.05 probability to win $90
	.85 probability to win $96

Participants were 420 undergraduates enrolled in lower division psychology; these were randomly assigned to either text or histogram conditions, with a .3 probability of assignment to text and .7 to histograms. We used a greater probability for assignment to the histogram condition because we expected the text (control) condition to replicate previous results and we wanted strong results for the new

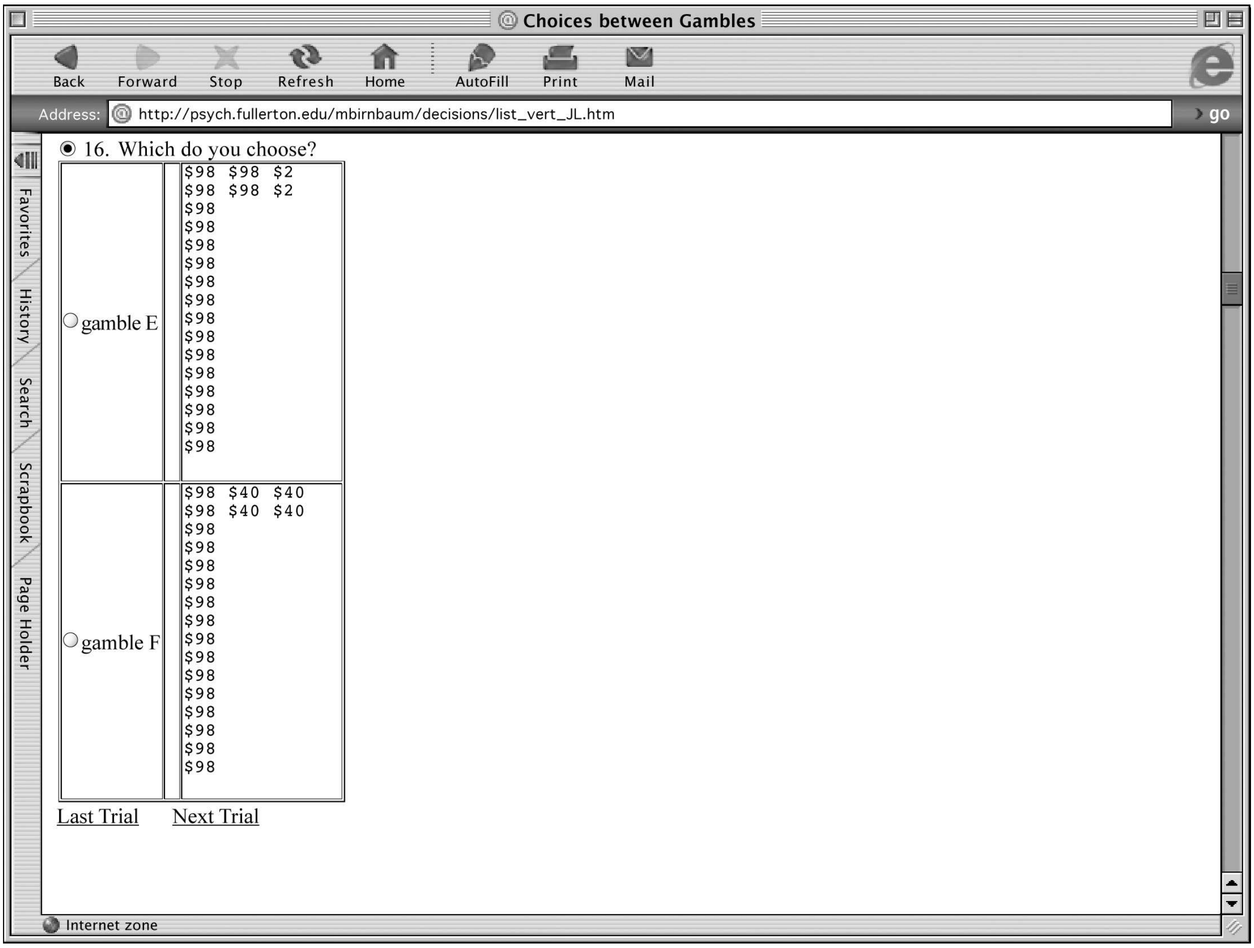

Figure 3: An example test of restricted branch independence in the vertical list format. If a person were to cancel the common consequences (in this case 16 tickets to win $98), then they would satisfy restricted branch independence.

condition. By random assignment, 125 participants received the *text* display, and 295 viewed *histograms*. The participants were 60% female and 91% were 22 years of age or younger.[1]

2.2 Study 2: Vertical Lists

The *list* format represented each gamble as an urn containing 20 equally likely tickets with prizes printed on them, as in Figures 2 and 3, which depict two of the choices. The values of each of the 20 tickets were displayed in vertical columns, instead of horizontal rows (as had been done in Birnbaum, 2004b). This way of presenting frequencies allows participants to "see" each equally likely consequence instead of seeing a number representing frequency. Each "branch" was a list presented in a separate column, which could be compared between gambles by a vertical eye movement. The translation of Birnbaum's (2004a) gambles to lists of 20 equally likely prizes required rounding in choices 7, 13, and 18 where .03 in probability was rounded to one ticket (0.05) in this study.

In the coalesced form of a choice, all consequences of a given value were printed in the same column (e.g., Figure 2). The split form of the same choice was created by placing the values for some of the consequences in a new column. In the canonical split form, (e.g., Figure 3), the number of tickets (probabilities) of corresponding ranked branches are equal and the number of branches is minimal. In this arrangement, it should be easy to identify common branches and cancel them, if a person wanted to do so. Each participant received all 20 choices, which included both coalesced and split forms of the key choices.

[1] Materials can be examined via the following URL: `http://psych.fullerton.edu/mbirnbaum/decisions/kj_histograms/conditionassignment.htm`
The two conditions can be examined at the following URLs: `http://psych.fullerton.edu/mbirnbaum/decisions/kj_histograms/graphcondition.htm`; `http://psych.fullerton.edu/mbirnbaum/decisions/kj_histograms/textcondition.htm`.

Table 2: Violations of stochastic dominance and coalescing. Table entries are percentages of violation of stochastic dominance in Birnbaum's (2004b) "Tickets" condition, the Text and Histograms conditions of Study 1, and in the Vertical list format of Study 2. Entries in bold show where significantly more than half of participants violated stochastic dominance as predicted by the prior TAX model (Appendix B).

	Choice		Condition			
No.	G^+	G^-	Tickets (342)	Text (125)	Histogram (295)	Vertical (408)
5	0.90 to win $96 0.05 to win $14 0.05 to win $12	0.85 to win $96 0.05 to win $90 0.10 to win $12	**71**	**70**	**72**	**81**
11	0.85 to win $96 0.05 to win $96 0.05 to win $14 0.05 to win $12	0.85 to win $96 0.05 to win $90 0.05 to win $12 0.05 to win $12	06	21	16	12
7	0.94 to win $99 0.03 to win $8 0.03 to win $6	0.91 to win $99 0.03 to win $96 0.06 to win $6	**67**	**61**	**76**	**86**
13	0.91 to win $99 0.03 to win $99 0.03 to win $8 0.03 to win $6	0.91 to win $99 0.03 to win $96 0.03 to win $6 0.03 to win $6	13	13	11	16

Notes: The dominant gamble (G^+) was presented first in choices 5 and 11 and second in choices 7 and 13. All choice percentages in the table are significantly different from 50%. Note that choices 5 and 11 are the same, except for coalescing, as are choices 7 and 13.

2.3 Design of Study 2: Coalescing and Branch Independence

Choices for Series A and B of Allias paradoxes are shown in Tables 5 and 6, respectively. Each choice is created from the choice directly above it in the tables by either coalescing (splitting) or by restricted branch independence. In Series A, the common branch is either 16 tickets to win $2 (first two rows), 16 tickets to win $40 (middle row), or 16 tickets to win $98 (last two rows). In Series B, the common branch has 17 tickets to win $7 (first two rows), $50 (third row), or $100 (fourth and fifth rows). Positions (First or Second) of *S* (the "safe" gamble with higher probabilities to win a smaller prize) and *R* ("risky" gamble) are counterbalanced between Series A and B.[2]

Participants were 408 undergraduates from the same subject pool as in the first study (61% female and 91% were 22 or younger). The two studies were embedded among a dozen other studies of judgment and decision making. They were separated by at least two other tasks that required an intervening time of 10 minutes or more.

[2]Complete materials can be viewed at URL: `http://psych.fullerton.edu/mbirnbaum/decisions/List_vert_JL.htm`.

3 Results

Tests of stochastic dominance and coalescing are presented in Table 2 for both studies. The percentage of violations of stochastic dominance is shown in the table for each condition of Study 1 (Text and Histogram) and Study 2 (Vertical list). The results from the "Tickets" condition of Birnbaum (2004b) with 342 participants are shown for comparison. Results show that the two new formats, histogram and vertical list appear to give rates of violation that are about the same or even higher than those in the text (control group) of Study 1 (labeled "Text") or in Birnbaum's (2004b) tickets condition.

Choice 11 is the same as Choice 5, except that it is presented in canonical split form; similarly Choices 7 and 13 differ only in form. According to CPT, people should make the same decision in Choices 5 and 11; and they should reach the same decision in Choices 7 and 13. However, the data in all formats show that significantly more than half of all participants violated stochas-

Table 3: Tests of stochastic dominance and coalescing in choices 5 and 11. Each entry shows the number of choices of each pattern in the different studies. (Row totals may not equal the number of participants, due to occasional skipping of an item.). The percentages of the $G^{-}GS^{+}$ preference pattern (predicted by TAX) are 65%, 47%, 61%, and 72% for the Tickets condition of Birnbaum (2004b), Text, Histograms, and Vertical List conditions.

	Choice Pattern			
Condition	$G^{+}GS^{+}$	$G^{+}GS^{-}$	$G^{-}GS^{+}$	$G^{-}GS^{-}$
Tickets (342)	95	3	**224**	16
Text (125)	17	21	**59**	28
Histograms (295)	43	38	**179**	34
Vertical List (408)	63	14	**294**	35

tic dominance in coalesced form, and significantly fewer than half of all participants violated stochastic dominance when the same choices are presented in canonical split form. (In canonical split form of a choice, both gambles are split so that there are equal probabilities on corresponding ranked branches and the number of branches is equal in both gambles and minimal). This large change in choice proportions represents a strong violation of coalescing.

In order to examine the extent to which aggregate results are representative of individuals' preference patterns, we counted the number of participants who made each pattern of responses. Table 3 shows the number of participants who showed each combination of choices for Problems 5 and 11 in all three new conditions, with previous data from the "Tickets" condition of Birnbaum (2004b) included for comparison.

In all cases, the modal (most frequent) response combination is to violate stochastic dominance in coalesced form and to satisfy it in the canonical split form ($G^{-}GS^{+}$). The relative frequency of this pattern (which contradicts CPT and is predicted by TAX) is even higher in the histogram and vertical conditions than it is in the text (control) condition. Significantly more than half of the sample in each of the two new conditions (61% and 72% in histogram and vertical list conditions, respectively) show the preference combination, $G^{-}GS^{+}$, compared with 47% in the control (replication) condition. Comparing text and histograms conditions (the two conditions—rows in Table 3—to which participants were randomly assigned), this difference is significant, $\chi^2(3) = 10.9$, $p < .05$.

Similar results were observed for Choices 7 and 13. Despite appealing intuitions that these formats would reduce violations of stochastic dominance and coalescing, there is no evidence that they did so. Instead, the new conditions produced stronger refutation of CPT. (Additional analyses of individual preference patterns, as in Table 3, are presented in Appendix F for other properties tested).

Table 4 shows an analysis of upper and lower cumulative independence and restricted branch independence in Study 1. The choice proportions in the two new conditions are quite similar to those of previous results (data from the "Tickets" condition of Birnbaum (2004b) are included in Table 4 for comparison). Choice proportions are quite similar in the two conditions of Study 1. Across all 20 choices, the correlation between choice proportions in the text (control) and the histogram format is 0.97; similarly, the correlation between the previous "Tickets" results and the new histogram proportions is 0.95.

Choices 10 and 9 represent a test of upper cumulative independence, which is implied by CPT. According to this property, the percentage choosing the risky gamble in Choice 9 should exceed that of Choice 10. Instead, the percentages choosing the risky gamble are significantly greater in Choice 10 than in Choice 9 in all three sets of data. According to PH, people should choose the "safe" gamble in both Choices 10 and 9 (higher worst consequence), but significantly more than half does the opposite in Choice 10. These results thus contradict both CPT and PH. Appendix F shows that significantly more people changed preferences in violation of the property than changed preferences consistent with it. Choices 12 and 14 represent another test of upper cumulative independence. Again, results contradict the property, which is implied by CPT with any functions and parameters.

Choices 6 and 8 test lower cumulative independence. According to this property, the proportion choosing the "safe" gamble in Choice 8 should be at least as great as that in Choice 6. Instead, the results show the opposite pattern. According to the priority heuristic, most people should have chosen the "safe" gamble in Choice 8. Instead, most people did the opposite. Choices 17 and 20 provide another test of lower cumulative independence; again, the percentage choosing the "safe" gamble is significantly lower in Choice 20 than in Choice 10, contradicting CPT with any probability weighting and utility functions. See Appendix F.

Choices 6 and 10 show a test of restricted branch independence, as do Choices 17 and 12. In both cases, more people who showed the pattern, $S \succ R$ and $R' \succ S'$ than showed the opposite reversal. This pattern contradicts the CPT model with an inverse-S decumulative weighting function but is consistent with the TAX model (Appendices B and F).

Tables 5 and 6 show a dissection of the Allais paradoxes into restricted branch independence and coalescing in Series A and B, respectively. For $n = 408$, choice

Table 4: Tests of upper cumulative independence (Choices 10, 9, 12, and 14) and tests of lower cumulative independence (Choices 6, 8, 17, and 20) in Study 1. The "Tickets" condition shows previous results from Birnbaum (2004b). Entries in the last three columns are percentages of choices of the "risky" gamble.

		Choice		Condition		
No.	Type	Safe Gamble, S	Risky Gamble, R	Tickets (342)	Text (125)	Histogram (295)
10	S' *vs.* R'	($110, 0.8; $44, 0.1; $40, 0.1)	($110, 0.8; $98, 0.1; $10, 0.1)	74	74	67
9	S''' *vs.* R'''	($98, 0.8; $40, 0.2)	($98, 0.9; $10, 0.1)	33	46	37
12	S' *vs.* R'	($106, 0.9; $52, .05; $48, .05)	($106, 0.9; $96, .05; $12, .05)	54	57	53
14	S''' *vs.* R'''	($96, 0.9; $48, 0.1)	($96, 0.95; $12, .05)	22	19	17
6	S *vs.* R	($44, 0.1; $40, 0.1; $2, 0.8)	($98, 0.1; $10, 0.1; $2, 0.8)	63	51	54
8	S'' *vs.* R''	($44, 0.2; $10, 0.8)	($98, 0.1; $10, 0.9)	70	70	75
17	S *vs.* R	($52, .05; $48, .05; $3, 0.9)	($96, .05; $12, .05; $3, 0.9)	46	44	42
20	S'' *vs.* R''	($52, 0.1; $12, 0.9)	($96, .05; $12, 0.95)	69	86	94

Note: In Choices 10, 9, 6, and 8, the "Safe" Gamble was presented first; in Choices 12, 14, 17, and 20 it was presented second.

percentages outside the interval from 45.1% to 54.9% are significantly different from 50% by two-tailed test with $\alpha = 0.05$. Data from Birnbaum (2004a, averaged over all 350 participants) are included in Tables 5 and 6 for comparison. The last four columns in Tables 5 and 6 show predictions of TAX and CPT using "prior" parameters (Appendices A and B). TAX and CPT disagree only in Choices 9 and 16 of Table 5 and Choices 17 and 14 of Table 6.

Each choice problem in Table 5 differs from the one in the row above by either restricted branch independence or by coalescing. For example, the first two choices in Table 5 (Choice Problems #6 and 9) differ only by coalescing. In Choice Problem 9, the safe gamble has two branches of 2 tickets to win $40 and a branch of 16 tickets to win $2. These two branches to win $40 have been coalesced in Choice 6. Similarly, the risky gamble in Choice Problem 9 has two branches of 2 and 16 tickets each to win $2. These have been coalesced in Choice 6. According to CPT (with any functions and parameters), people should make the same decisions in these two choices. However, note that whereas 62% chose the risky gamble in Choice 6, only 45% did so in Choice 9 in the vertical list condition. This difference is significant by the test of correlated proportions, $z = 5.67$.

Choices 9, 12, and 16 differ only in the consequence on the common branch of 16 tickets; these consequences were $2, $40, and $98 in Choices 9, 12, and 16, respectively. If people cancelled common branches, they would make the same decisions in all three cases. Instead, we see that more than half chose the "safe" gamble in Choice 9 whereas more than half chose the "risky" gamble in Choice 16. This trend agrees with that in Birnbaum (2004a) and is also significant with the vertical list, $z = 3.20$.

Finally, note that Choice Problems #16 and 19 differ in that the two branches to win $98 in the risky gamble have been coalesced in #19, as have the two branches to win $40 in the safe gamble. The choice percentage changes from 55% to 11%, $z = 13.50$, contrary to CPT.

Similar results were obtained in Table 6 (Series B), where the (first/second) positions of the risky and safe gambles have been counterbalanced. The modal choices resemble those of Table 5 and of Birnbaum (2004a), except in Choice #17, where the choice percentage was 73% in the new study compared to 48% in the previous study. The test of coalescing between Choices 10 and 17 showed much smaller, but still significant effects, $z = 2.25$. The test of restricted branch independence of Choices 17 and 14 fell short of significance, $z = 1.93$. As in Series A and the previous study, violations of restricted branch independence in Choices 20 and 14 were significant, $z = 9.02$, as were the violations of coalescing in Choices 14 and 8 (78% versus 29%; $z = 13.19$). See Appendix F.

The results show that the new data with vertical lists show violations of restricted branch independence, which contradict original prospect theory and which are opposite the pattern predicted by CPT with an inverse-S weighting function. In addition, there are systematic violations of coalescing, which contradict CPT with any weighting function. The modal choices in the vertical list condition agree with previous results and with pre-

Table 5: Dissection of Allais Paradox (Series A). Each entry under "Condition" is the percentage choosing the risky gamble, which was presented first in Series A. Data from Birnbaum (2004a) are aggregated over all 350 participants in that study. Last four columns show predicted certainty equivalents of the gambles according to TAX and CPT using prior parameters.

		Choice		Condition		Prior TAX		Prior CPT	
No.	Relation to previous row	"Safe" gamble, *S*	"Risky" gamble, *R*	Birnbaum (2004a)	Vertical List	*S*	*R*	*S*	*R*
6		4 tickets to win $40 16 tickets to win $2	2 tickets to win $98 18 tickets to win $2	62	62	9.0	13.3	10.7	16.9
9	Split # 6	2 tickets to win $40 2 tickets to win $40 16 tickets to win $2	2 tickets to win $98 2 tickets to win $2 16 tickets to win $2	36	45	11.1	9.6	10.7	16.9
12	RBI #9	2 tickets to win $40 16 tickets to win $40 2 tickets to win $40	2 tickets to win $98 16 tickets to win $40 2 tickets to win $2	46	43	40.0	30.6	40.0	38.0
16	RBI #9, 12	16 tickets to win $98 2 tickets to win $40 2 tickets to win $40	16 tickets to win $98 2 tickets to win $98 2 tickets to win $2	57	55	59.8	62.6	74.5	67.6
19	Coalesce #16	16 tickets to win $98 4 tickets to win $40	18 tickets to win $98 2 tickets to win $2	22	11	68.0	54.7	74.5	67.6

Notes: The common branch was 16 tickets to win $2 in Choices 6 and 9, $40, in Choice 12, and $98 in Choices 16 and 19, respectively. RBI = Restricted Branch Independence, *R* = Risky Gamble, *S* = Safe Gamble.

dictions of the "prior" TAX model in Tables 2, 5 and 6 (except Choice 17).

The priority heuristic implies that most people should have chosen the safe gamble in Choices 6 and 16 and the risky gamble in Choice 9. Instead, the majority chose the risky gamble in Choices 6 and 16 and the safe gamble in Choice 9. It also predicts that the majority should have chosen the safe gamble in Choice 14 of Table 6, whereas most people chose the risky gamble on that choice.

4 Discussion

These results were quite surprising to two authors of this paper, who were confident that the graphical format for representing relative frequency and especially the combined graphical and list format of the second study would produce results more consistent with principles that had been violated in previous research. In particular, it was conjectured that the methods used here would produce data consistent with stochastic dominance, coalescing, and restricted branch independence.

Instead, the present results yield similar findings to those of previous research (Birnbaum, 2004a; 2004b; 2006; 2007) using a dozen other formats. The new results are consistent with the interpretation that previous failures of CPT, EU, and PH are not due to specific details of display procedures, but rather are generated instead by deeper processes for evaluation and choice. These results should be comforting to theorists because they suggest that basic findings in judgment and decision making can indeed be replicated using different ways of presenting choices. One need not design a new theory for each new format.

The findings violate CPT, which fails to correctly predict violations of stochastic dominance, the violations of coalescing, and which predicts the opposite pattern of violation of restricted branch independence from what is observed. In addition, results contradict predictions of the PH. Brandstätter et al. (2008) have acknowledged these problems and suggested that perhaps people use different heuristics when presented with different choices. For criticism of that approach, see Birnbaum (2008a, 2008c).

Consistent with previous research, majority violations of stochastic dominance are found with both of the new formats used here (Table 2). Table 2 also shows that these violations can be strongly reduced and nearly eliminated by presenting choices in canonical split form (in which probabilities of corresponding ranked branches are equal and the number of branches in the gambles are minimal). These findings, summarized in Table 3, violate both CPT (which implies coalescing and stochastic dominance) and

Table 6: Dissection of Allais paradox (Series B), as in Table 5. In Series B, the "safe" and "risky" gambles were presented first and second, respectively, counterbalancing the arrangement of Series A. Entries show percentages choosing the "risky" gamble.

No.	Relation to previous row	Choice: "Safe" gamble, *S*	Choice: "Risky" gamble, *R*	Condition: Birnbaum (2004a)	Condition: Vertical List	Prior TAX: *S*	Prior TAX: *R*	Prior CPT: *S*	Prior CPT: *R*
10		3 tickets to win $50 17 tickets to win $7	2 tickets to win $100 18 tickets to win $7	81	78	13.6	18.0	15.9	22.1
17	Split #10	2 tickets to win $50 1 ticket to win $50 17 tickets to win $7	2 tickets to win $100 1 tickets to win $7 17 tickets to win $7	48	73	15.6	14.6	15.9	22.1
20	RBI #17	2 tickets to win $50 17 tickets to win $50 1 ticket to win $50	2 tickets to win $100 17 tickets to win $50 1 ticket to win $7	49	50	50.0	40.1	50.0	49.2
14	RBI # 17, 20	17 tickets to win $100 2 tickets to win $50 1 ticket to win $50	17 tickets to win $100 2 tickets to win $100 1 ticket to win $7	62	78	68.4	69.7	82.2	79.0
8	Coalesce #14	17 tickets to win $100 4 tickets to win $50	19 tickets to win $100 1 ticket to win $7	26	29	75.7	62.0	82.2	79.0

original prospect theory (which assumed that people use the editing rule of combination, which would imply coalescing in these studies).

Stripped prospect theory (subjectively weighted utility without the editing rules) implies restricted branch independence (Appendix E). Violations of restricted branch independence in Studies 1 and 2 contradict predictions of that stripped version of original prospect theory.

Study 1 replicated violations of lower and upper cumulative independence with the text format and found that these violations persist with the histogram format (Table 4). These violations violate EU and CPT with any utility and probability weighting functions.

Study 2 shows that Allais paradoxes are replicated with the vertical list format. Contrary to the theory that people would cancel common branches (tickets that are the same in both urns) in the vertical list format, Study 2 found that both coalescing and restricted branch independence are violated (Tables 5 and 6). Study 2 also replicated previous results quite closely, except for Choice 17 of Table 6.

Because of the many manipulations (reviewed in our introduction) that do affect preference orders, and because of previous failures to replicate weak effects with studies lacking power, some investigators have cautioned that new results should not be taken seriously until they have been replicated (e.g., Evanschitzky, Baumgarth, Hubbard, & Armstrong, 2007).[3]

In relation to previous failures to replicate, four features of the present research should perhaps be noted. First, the phenomena selected for replication are large effects; for example, the violations of stochastic dominance in the coalesced form are often 70% or greater among undergraduates and 20% or less in canonical split form. Second, we tested enough participants in order to achieve sufficient power to make accurate estimates of the incidence of these phenomena. Third, we included a replication study using text format with random assignment of the new participants to conditions, in order to check that the original findings can be replicated and thus permit an experimental comparison of the format manipulation of Study 1. Fourth, we tested participants who were from the same "subject pool" as used in previous research. When studies lack these features, failures to replicate are difficult to interpret.

With respect to the question of what is the "best" way to display probabilistic information in applied decision making, as in medical decision making (Lipkus, 2007; Cutie, et al. 2008), one result is quite clear: If we think that people should satisfy stochastic dominance (which most theorists agree is normative), we should present choices in canonical split form, whether this is displayed by text, histograms, or lists (Table 2). In all formats, this manipulation (form) strongly reduces violations of stochastic dominance. When Allais paradoxes are presented in split form, however, they are not merely re-

[3]See the discussions in `http://www.sjdm.org/mail-archive/jdm-society/2007-May/002977.html`.

duced, they are actually reversed (Table 5). Therefore, people do not conform to EU theory, simply because choices are presented in canonical split form. Restricted branch independence is a stronger form of Savage's "sure thing" axiom, an axiom that has been questioned not only as a descriptive principle but also as a normative principle (Allais & Hagen, 1979). So if one were to recommend that all choices be presented in canonical split form, one would be endorsing a theory that satisfies dominance and violates restricted branch independence, but the violations of restricted branch independence people exhibit are ironically opposite what had been intuited to be reasonable.

Contrary to expectations of two coauthors, we were unable to find procedures that would yield data compatible with CPT or the PH. These studies confirmed previous findings and extended them to a wider domain.

Appendix A: CPT

Cumulative prospect theory (CPT) has the same representation as rank and sign-dependent utility theory (Tversky & Kahneman, 1992; Luce & Fishburn, 1991, 1995). For gambles consisting of strictly positive consequences, $G = (x_1, p_1; x_2, p_2; \ldots; x_n, p_n)$, $x_1 \geq x_2 \geq \ldots \geq x_n > 0$, $p_1 + p_2 + \ldots + p_n = 1$, this model can be written as follows:

$$U(\mathrm{G}) = \sum \mathrm{w}_i u(x_i) \qquad (\mathrm{A.1})$$

where $\mathrm{w}_i = W(P_i) - W(P_{i-1})$ where $P_i = p_1 + p_2 + \ldots + p_i$ and $p_0 = 0$; the function, W (the decumulative probability weighting function), is strictly monotonic such that $W(0) = 0$ and $W(1) = 1$. As shown in Birnbaum and Navarrete (1998), this theory satisfies first order stochastic dominance, coalescing, and both lower and upper cumulative independence (first four properties in Table 1). The properties of lower and upper cumulative independence were devised by Birnbaum (1997) in order to provide tests between his configural weight models and rank dependent models like CPT.

The CPT model has been fit to data with the assumptions that $u(x) = x^{\beta}$ and with the following probability weighting function (Tversky & Fox, 1995):

$$W(P) = \frac{cP^{\gamma}}{cP^{\gamma} + (1 - P)^{\gamma}} \qquad (A.2)$$

When $\gamma < 1$, this function has an inverse-S shape, showing smaller changes when P is near 0.5 and relatively greater changes near 0 and 1. Predictions calculated using parameter values previously published in the literature ($\beta = 0.88$, $c = 0.724$; $\gamma = 0.61$) are called predictions of the "prior" CPT model.

For the examples in the right column of Table 1, the computed utilities of the gambles in the "prior" CPT model are $U(G^{+}) = 72.3 > U(G^{-}) = 71.7$; $U(GS^{+}) = 72.3 > U(GS^{-}) = 71.7$; $U(S) = 10.4 < U(R) = 15.3$; $U(S'') = 17.5 < U(R'') = 22.3$; $U(S') = 83.5 > U(R') = 80.1$; $U(S''') = 75.7 > U(R''') = 72.8$. With these parameters, CPT predicts $R \succ S$ and $S' \succ R'$ for the choices in the last row and column of Table 1. As shown in Birnbaum (2008b), the CPT model with any inverse-S weighting function implies that if violations of restricted branch independence are observed, they will be of this pattern (RS'), which is opposite the pattern predicted by TAX, described in Appendix B. Instead, data show that the SR' combination of preferences is significantly more frequent than RS'. This model is incorrect in predicting the modal responses in Choices 5 and 7 of Table 2, of Choices 10, 12, and 17 of Table 4, of #9 and 16 of Table 5, and of Choice 14 of Table 6.

Appendix B: The TAX model

Birnbaum's (1999a) transfer of attention exchange (TAX) model assumes that the utility of a gamble is a weighted average of the utilities of the consequences in which the weights depend on the ranks of the discrete branches and their probabilities. The "special TAX model" can be written as follows for two-branch gambles, $G = (x_1, p_1; x_2, p_2)$, where $x_1 \geq x_2 > 0$, $p_1 + p_2 = 1$, when $\delta > 0$,

$$U(G) = \frac{au(x_1) + bu(x_2)}{a + b} \qquad (B.1)$$

where $a = t(p_1) - \delta t(p_1)/3$
$b = t(p_2) + \delta t(p_1)/3$

Eq. 1 is a weighted average of consequence utilities, where weights depend on probabilities of the consequences and ranks of the consequences on discrete branches. The expression, $\delta/3$, represents the proportion of probability weight, $t(p)$, transferred from the branch with the higher consequence to the branch with the lower consequence.

For three branch gambles, $G = (x_1, p_1; x_2, p_2; x_3, p_3)$, where $x_1 \geq x_2 \geq x_3 > 0$, $p_1 + p_2 + p_3 = 1$, and $\delta > 0$, the model is as follows:

$$U(G) = \frac{Au(x_1) + Bu(x_2) + Cu(x_3)}{A + B + C} \qquad (B.2)$$

where $A = t(p_1) - 2\delta t(p_1)/4$
$B = t(p_2) - \delta t(p_2)/4 + \delta t(p_1)/4$
$C = t(p_3) + \delta t(p_1)/4 + \delta t(p_2)/4$

In practice, the weighting function is approximated by $t(p) = p^{\gamma}$, where $0 < \gamma \leq 1$, and $u(x) = x^{\beta}$, where $0 < \beta \leq 1$. When the model is fit to individual data, Birnbaum and Chavez (1997) reported a median estimated value of $\beta = 0.61$, whereas Birnbaum and Navarrete (1998) found a median estimate of β = 0.41. Many aggregated group data can be roughly approximated with $\beta = \delta = 1$ when consequences are small positive values of cash (e.g., Birnbaum, 1999a). With the assumptions that $\beta = \delta = 1$ and $\gamma = 0.7$, this model is called the "prior" TAX model, since these simplified parameters have been used to design experiments and make predictions to group data for new studies (e.g., Birnbaum, 1999a, 2004a, 2006, 2007). [This model is the same as that in Birnbaum (1999a), except a notational convention has been changed so that $\delta > 0$ here corresponds to $\delta < 0$ in Birnbaum (1999a) and earlier papers.]

This "prior" TAX model predicts the pattern of choices shown in the right column Table 1. The computed utilities of these gambles are $U(G^+) = 45.7 < U(G^-) = 63.1$; $U(GS^+) = 53.1 > U(GS^-) = 51.4$; $U(S) = 11.4 > U(R) = 10.9$; $U(S'') = 16.2 < U(R'') = 20.4$; $U(S') = 65.0 < U(R') = 69.6$; $U(S''') = 68.0 > U(R''') = 58.3$. This prior model correctly predicted all but two of the modal choices in Tables 1−6 (Choice 6 in Table 4 and Choice 17 in Table 6).

Appendix C: Priority heuristic

According to the priority heuristic (PH) of Brandstätter, et al. (2006), people first compare lowest consequences of a gamble and choose the gamble with the higher lowest consequence if this difference exceeds 10% of the largest consequence in either gamble, rounded to the nearest prominent number ($10 in choices of this paper). When the lowest consequences differ by less (than $10), people supposedly choose the gamble with the smaller probability to get the lowest consequence, if these differ by 0.1 or more.

If the probabilities of the lowest consequences differ by less than 0.1, the person is theorized to next compare the highest prizes and choose by that criterion. When there are more than two branches and the first three comparisons yield no decision, people next compare the probabilities to win the highest prize and decide on that basis alone, if there is any difference. If all four criteria yield no choice, the person chooses randomly. When gambles have three or more branches, the PH assumes that people never examine intermediate branches.

With respect to the examples in Table 1, the PH predicts that people should satisfy stochastic dominance in the first choice because the lowest consequences are the same and the probabilities of these consequences differ by less than 0.1; the highest consequences are the same, but G^+ has the higher probability to win the best consequence. In the second choice, all four features that are considered by PH (lowest and highest consequences and their probabilities) are the same, so the person should be indifferent between these two gambles. In the third row of Table 1 (testing lower cumulative independence), a person should choose R over S (because of the higher best consequence) and S'' over R'' (because of the probability of the lowest consequence). In the test of upper cumulative independence, a person should choose S' over R' (lowest consequence) and S''' over R''' (lowest consequence). In the tests of restricted branch independence (last row of Table 1), a person should choose R over S (highest consequence) and S' over R' (lowest consequence). Thus, the PH model agrees with the examples in the right most column of Table 1 only in the choice of S''' over R''' in the fourth row. The PH does not predict any of the four modal choices in Table 2, is incorrect in Choices 10, 12, 8, and 17 of Table 4, and fails to predict results of Choices 6, 9, 14, and 16 of Tables 5 and 6.

Appendix D: Expected-utility theory and Allais paradoxes

According to EU theory, $S = (y, p; z, 1 - p) \succ R = (x, q; z, 1 - q) \Leftrightarrow S'' = (y, 1) \succ R'' = (x, q, y, 1 - p; z, p - q) \Leftrightarrow S' = (x, 1 - p; y, p) \succ R' = (x, 1 - p + q; z, p - q)$. It is assumed that $x > y > z \geq 0$, $1 \geq p > q \geq 0$, and that all of the probabilities are between 0 and 1. Proof: According to Expected Utility theory, $S = (y, p; z, 1 - p) \succ R = (x, q; z, 1 - q) \Leftrightarrow pu(y) + (1 - p)u(z) > qu(x) + (1 - q)u(z) \Leftrightarrow pu(y) + (1 - p)u(z) > qu(x) + (p - q)u(z) + (1 - p)u(z)$. There is a common branch of $(z, 1 - p)$ in both gambles. We can change the value of z on this branch in both gambles without reversing the preference order. If we change z to y in both gambles, for example, we have $pu(y) + (1 - p)u(y) = u(y) > qu(x) + (p - q)u(z) + (1 - p)u(y) \Leftrightarrow S'' = (y, 1) \succ R'' = (x, q, y, 1 - p; z, p - q)$. The original form of Allais paradox (called the "type 1" paradox) is a special case of this choice between S and R compared with the choice between S'' and R''. If we change the consequence on the common branch from z to x, we have $\Leftrightarrow pu(y) + (1 - p)u(x) > qu(x) + (p - q)u(z) + (1 - p)u(x) \Leftrightarrow (1 - p)u(x) + pu(y) > (1 - \mathrm{p} + q)u(x) + (p - q)u(z) \Leftrightarrow S' = (x, 1 - p; y, p) \succ R' = (x, 1 - p + q; z, p - q)$. When people make different choices in S'' and R'' as opposed to S' and R', it is called the "type 2" paradox, and the conflict between S and R compared with the choice between S' and R' is called the "type 3" paradox.

More generally, it can be shown that any model that satisfies transitivity, coalescing, and restricted branch independence will not show these Allais paradoxes. Proof: From coalescing and transitivity, we have $S = (y, p; z, 1 - p) \succ R = (x, q; z, 1 - q) \Leftrightarrow (y, q; y, p - q; z, 1 - p) \succ (x, q; z, p - q, z, 1 - p)$. Note that there is a common branch of $(z, 1 - p)$ in both gambles. From restricted branch independence, we can change z to y in both gambles, so $\Leftrightarrow$ (y, q; y, $p - q$; y, $1 - p$) $\succ$ (x, q; z, $p - q$, y, $1 - p$). Using coalescing and transitivity, this is equivalent to $S'' \succ R''$. Similarly, by restricted branch independence we can change z to x in the common branch, so (x, $1 - p$; y, q; y, $p - q$;) $\succ$ (x, $1 - p$; x, q; z, $p - q$). By coalescing and transitivity, we see that this is equivalent to $S' \succ R'$. Therefore, any transitive theory must violate either branch independence or coalescing to account for these Allais paradoxes.

Thus, Allais paradoxes can be described as a confounded test of coalescing and restricted branch independence. To compare original prospect theory, CPT and TAX, we must dissect the Allais paradoxes (as in Tables 5 and 6) in order to separately test restricted branch independence and coalescing. Original prospect theory satisfies branch independence and violates coalescing (apart from its editing principle of combination) whereas CPT satisfies coalescing and violates branch independence (apart from the editing principle of cancellation). The TAX model violates both properties but implies that the Allais paradoxes are due to violations of coalescing and that violations of restricted branch independence actually reduce the magnitude of Allais paradoxes.

Appendix E: Stripped Prospect Theory

The following subjectively weighted utility model (Edwards, 1962) is sometimes called "stripped" prospect theory:

$$U(G) = \sum w(p_i)u(x_i)$$

where $G = (x_1, p_1; x_2, p_2; \ldots; x_n, p_n)$, and $\sum p_i = 1$. The term "stripped" is used to indicate that the editing rules and other special exceptions have been removed. This model satisfies restricted branch independence: $S = (x, p; y, q; z, 1 - p - q) \succ R = (x', p; y', q; z, 1 - p - q) \Leftrightarrow S' = (z', 1 - p - q; x, p; y, q) \succ R' = (z', 1 - p - q; x', p; y', q)$]. Proof: $S \succ R \Leftrightarrow$ U(x, p; y, q; z, $1 - p - q$) > U(x', p; y', q; z, $1 - p - q$) $\Leftrightarrow$ $w(p)u(x) + w(q)u(y) + w(1 - p - q)u(z) > w(p)u(x') + w(q)u(y') + w(1 - p - q)u(z)$. We can subtract the common term, $w(1 - p - q)u(z)$, from both sides and add the following to both sides, $w(1 - p - q)u(z')$, which holds $\Leftrightarrow S' \succ R'$. Therefore, systematic violations of restricted branch independence in Studies 1 and 2 contradict this "stripped" version of prospect theory, which uses no editing.

Table 7: Tests of Upper Cumulative Independence (Choices 10 and 9 in Study 1). S' = (\$110, 0.8; \$44, 0.1; \$40, 0.1) $\prec R'$ = (\$110, 0.8; \$98, 0.1; \$10, 0.1) $\Rightarrow$ S''' = (\$98, 0.8; \$40, 0.2) $\prec R'''$ = (\$98, 0.9; \$10, 0.1). The pattern, $R'S'''$, is inconsistent with this property, but is predicted by the TAX model with prior parameters.

	Choice Pattern			
Condition	$S'S'''$	$S'R'''$	$R'S'''$	$R'R'''$
Tickets (Birnbaum, 2004b)	71	16	**156**	96
Text	17	15	**49**	43
Histograms	73	24	**111**	86

Table 8: Tests of Lower Cumulative Independence in Choices 6 and 8 of Study 1: S = (\$44, 0.1; \$40, 0.1; \$2, 0.8) $\succ R$ = (\$98, 0.1; \$10, 0.1; \$2, 0.8) $\Rightarrow S''$ = (\$44, 0.2; \$10, 0.8) $\succ R''$ = (\$98, 0.1; \$10, 0.9). Violations predicted by TAX are shown in bold font.

	Choice Pattern			
Condition	SS''	SR''	RS''	RR''
Tickets (Birnbaum, 2004b)	52	**74**	51	163
Text	21	**40**	16	47
Histogram	49	**87**	25	134

Appendix F: Analysis of crosstabulation frequencies

Tables 7, 8, and 9 show the frequencies of choice combinations for tests of upper and lower cumulative independence and of restricted branch independence, for Choice Problems # 10, 9, 6, and 8 of Study 1. The new results with the histogram format are quite similar to previous results (Tickets condition from Birnbaum, 2004b) and to the results of the text (control) condition. The standard significance tests for systematic violations of the properties is the test of correlated proportions, which compares the frequency of predicted preference reversals against the frequency of the opposite type of reversal. All six of these tests (in Tables 7, 8, and 9 both the text and histograms conditions) are statistically significant, with values of z ranging from 3.6 to 7.5).

In Table 7, the $R'S'''$ response pattern (violations of upper cumulative independence) was shown by 46%, 39%, and 38% of the participants in tickets, text control and histograms condition. The distribution of response patterns was not significantly different between the text and histograms conditions, $\chi^2(3) = 7.3$. Similarly, in Table

Table 9: Tests of Restricted Branch Independence. Choices 6 and 10 (Study 1). The TAX model predicts the pattern, S = (\$44, 0.1; \$40, 0.1; \$2, 0.8) $\succ$ R = (\$98, 0.1; \$10, 0.1; \$2, 0.8) and S' = (\$110, 0.8; \$44, 0.1; \$40, 0.1) $\prec$ R' = (\$110, 0.8; \$98, 0.1; \$10, 0.1), shown in bold.

	Choice Pattern			
Condition	SS'	SR'	RS'	RR'
Tickets (Birnbaum, 2004b)	50	**77**	37	177
Text	19	**41**	13	50
Histograms	57	**79**	40	119

8, the rates of SR'' (violations of lower cumulative independence) are 22%, 32%, and 29%, respectively; the response patterns are not significantly different in the text and histograms conditions, $\chi^2(3)$ = 3.1. In Table 9, the rates of SR' are 23%, 33%, and 27% in previous data, text and histograms, respectively. The text and histograms conditions are not significantly different, $\chi^2(3)$ = 2.6. In sum, the violations of CPT in Study 1 are significant in all tests, and the magnitudes of violations (in 7, 8 and 9) are not significantly lower in the histograms condition compared with the text control.

Table 10 shows the numbers of participants who made each choice combination in the tests of the Allais paradoxes with the vertical list (Study 2). Both CPT and TAX make the same predictions for the three types of Allais paradoxes (Appendix D). The bold entries show that significantly more participants made the changes in preference predicted by these models than made the opposite reversals of preference. For example, in Choice 6 (Table 5), most people (62%) chose the risky gamble and in Choice 12 (Table 5) most people (89%) chose the safe gamble. The first row in Table 10 shows that 127 people showed this reversal of preference and 46 made the opposite reversal of preference. The second row shows similar results for the type 2 Allais paradox (Choices 12 and 19), and the third row shows that 221 people showed the predicted pattern of reversal for the type 3 paradox, compared to only 14 who had the opposite reversal of preference. The last three rows of Table 10 show that similar results were obtained for Series B (Table 6). All six tests are significant, with z ranging from 6.2 to 13.5.

Table 11 presents cross-tabulations that test coalescing. According to CPT, people should make the same decision in Choice 6 as they do in Choice 9, since these differ only in how branches are split or coalesced. The TAX model predicts that coalescing will be violated (predictions of the prior model are shown in bold). Table 11 shows that 106 people reversed preferences in the direction predicted

Table 10: Crosstabulations testing Allais paradoxes in Study 2. The entry under RS in the first row indicates that 127 people chose the "risky" gamble in Choice 6 and chose the "safe" gamble in Choice 12. This pattern, shown in bold, is consistent with typical results with the Allais paradox. Both CPT and TAX models with their prior parameters predict this pattern, shown in bold.

Allais Type	Choice Combination	RR	RS	SR	SS
1	6 × 12	127	**127**	46	106
2	12 × 19	27	**145**	18	215
3	6 × 19	31	**221**	14	139
1	10 × 20	172	**147**	34	54
2	20 × 8	77	**129**	42	160
3	10 × 8	98	**221**	21	67

Note: The first three rows show tests from Series A (Table 5) and the last three rows show results from Series B (Table 6).

Table 11: Crosstabulations testing coalescing (Study 2). Each entry shows the number of people who had each choice combination. Choices are specified in Tables 5 and 6. For example, the entry under RS in the first row shows that 106 people chose the "risky" gamble on Choice 6 (coalesced form) and the "safe" gamble on Choice 9 (split form) of Table 5.

Choice Combination	RR	RS	SR	SS
6 × 9	147	**106**	38	115
16 × 19	33	**191**	12	170
10 × 17	259	**59**	37	51
14 × 8	172	**147**	15	73

by the TAX model, and 38 had the opposite reversal of preference. Similarly, CPT predicts that people should make the same decision in Choice 16 as they do in Choice 19; instead 191 show the pattern of reversal predicted by TAX and only 12 show the opposite pattern. All four tests of coalescing are significant, with z ranging from 2.3 to 12.6. These results contradict CPT with any functions and parameters.

Table 12 shows the frequency of each choice combination in the tests of restricted branch independence (Study 2). All choices analyzed here were presented in the canonical split form in which the three corresponding ranked branches had equal probabilities. The common branch in Series A (Table 5) was either 16 tick-

Table 12: Crosstabulations showing tests of restricted branch independence in Study 2. These choices are all presented in canonical split form, in which common branches could be easily cancelled (see Figure 3). The entries in bold font show the patterns predicted by TAX model with prior parameters.

Choice Combination	*RR′*	*RS′*	*SR′*	*SS′*
9 × 12	97	*88*	75	**146**
12 × 16	122	50	**103**	131
9 × 16	130	55	**94**	128
17 × 20	164	*133*	41	**69**
20 × 14	184	22	**135**	66
17 × 14	248	49	**70**	39

ets to win \$2, 16 tickets to win \$40, or 16 tickets to win \$98 (Choices 9, 12, and 16, respectively). If a person were to cancel the common branch before choosing (which should be easy to do in the format of Figure 3), there would be no violations of restricted branch independence. Instead, both series yield similar conclusions with respect to the patterns of violations. When the common branch is moved from lowest to highest or from middle to highest, there are more people who switch from the safe gamble to the risky gamble (*SR′*), and the opposite pattern is more frequent when the common consequence is shifted from lowest to middle ranked branch (italics). The CPT model predicts violations of restricted branch independence. However, the inverse-S weighting function implies the opposite pattern of violations from the most frequently observed pattern (it predicts *RS′* in the cases where the *SR′* pattern is more frequent. Note also that the pattern of violations of branch independence (9 × 16) is opposite the direction of the Allais paradox (6 × 19). According to CPT, these should have been in the same direction. A similar pattern of results is observed for Series B (Table 12). The statistical tests yielded values of $z = -1.0$, 4.3, 3.2, -7.0, 9.0, and 1.9 for the six rows of Table 12, respectively (values of $|z| > 2$ are "significant" with $p < .05$). Note in Tables 5 and 6 that the difference in TAX favoring the safe gamble is much smaller in Choice 9 than in 12 and smaller in Choice 17 than in 20; such a difference might result in greater relative frequency of the *RS′* pattern, shown in Italics for 9 × 12 and 17 × 20.

In summary, these analyses confirm that the patterns observed in the choice proportions are also characteristic of individual data and that patterns of violations of CPT predicted by the TAX model are statistically significant.

References

Allais, M., & Hagen, O. (Ed.). (1979). *Expected utility hypothesis and the Allais paradox*. Dordrecht, The Netherlands: Reidel.

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

Birnbaum, M. H. (1999a). Paradoxes of Allais, stochastic dominance, and decision weights. In J. Shanteau, B. A. Mellers, & D. A. Schum (Eds.), *Decision science and technology: Reflections on the contributions of Ward Edwards* (pp. 27–52). Norwell, MA: Kluwer Academic Publishers.

Birnbaum, M. H. (1999b). Testing critical properties of decision making on the Internet. *Psychological Science, 10*, 399–407.

Birnbaum, M. H. (2001). A Web-based program of research on decision making. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 23–55). Lengerich, Germany: Pabst Science Publishers.

Birnbaum, M. H. (2004a). Causes of Allais common consequence paradoxes: An experimental dissection. *Journal of Mathematical Psychology, 48*, 87–106.

Birnbaum, M. H. (2004b). Tests of rank-dependent utility and cumulative prospect theory in gambles represented by natural frequencies: Effects of format, event framing, and branch splitting. *Organizational Behavior and Human Decision Processes, 95*, 40–65.

Birnbaum, M. H. (2005). Three new tests of independence that differentiate models of risky decision making. *Management Science, 51*, 1346–1358.

Birnbaum, M. H. (2006). Evidence against prospect theories in gambles with positive, negative, and mixed consequences. *Journal of Economic Psychology, 27*, 737–761.

Birnbaum, M. H. (2007). Tests of branch splitting and branch-splitting independence in Allais paradoxes with positive and mixed consequences. *Organizational Behavior and Human Decision Processes, 102*, 154–173.

Birnbaum, M. H. (2008a). Evaluation of the priority heuristic as a descriptive model of risky decision making: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review, 115*, 253–262.

Birnbaum, M. H. (2008b). New paradoxes of risky decision making. *Psychological Review, 115*, 463–501.

Birnbaum, M. H. (2008c). New tests of cumulative prospect theory and the priority heuristic: Probability-outcome tradeoff with branch splitting. *Judgment and Decision Making, 3*, 304–316.

Birnbaum, M. H., & Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes, 71*, 161–194.

Birnbaum, M. H., & Navarrete, J. B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty, 17*, 49–78.

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

Birnbaum, M. H., & Zimmermann, J. M. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. *Organizational Behavior and Human Decision Processes, 74*, 145–187.

Bleaney, M., & Humphrey, S. J. (2006). An experimental test of generalized ambiguity aversion using lottery pricing tasks. *Theory and Decision, 60*, 257–282.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Choices without tradeoffs. *Psychological Review, 113*, 409–432.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky Choice with Heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, & Willemsen (2008) and Rieger & Wang (2008). *Psychological Review, 115*, 281–289.

Camerer, C. F. (1998). Bounded rationality in individual decision making. *Experimental Economics, 1*, 163–183.

Cutie, C. L., Weinstein, N. D., Emmons, K., & Colditz, G. (2008). A test of numeric formats for communicating risk probabilities. *Medical Decision Making, 28*, 377–384.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review, 69*, 109–135.

Evanschitzky, H., Baumgarth, C., Hubbard, R., & Armstrong, J. S. (2007). Replication research's disturbing trend. *Journal of Business Research, 60*, 411–415.

Fox, C. R., & Hadar, L. (2006). "Decisions from experience" = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making, 1*, 159–161.

Goldstein, W., & Einhorn, H. J. (1987). A theory of preference reversals. *Psychological Review, 94*, 236–242.

Harless, D. W. (1992). Actions versus prospects: The effect of problem representation on regret. *The American Economic Review, 82*, 634–649.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choices. *Psychological Science, 15*, 534–539.

Huff, D. (1954). *How to lie with statistics*. New York: Norton.

Humphrey, S. J. (1995). Regret aversion or event-splitting effects? More evidence under risk and uncertainty. *Journal of risk and uncertainty, 11*, 263–274.

Kahneman, D. (2003). Experiences of collaborative research. *American Psychologist, 58*, 723–730.

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

Keller, L. R. (1985). The effects of problem representation on the sure-thing and substitution principles. *Management Science, 31*, 738–751.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology of framing effects. *Organizational Behavior and Human Decision Processes, 76*, 149–188.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*, 46–55.

Lindman, H. R. (1971). Inconsistent preferences among gambles. *Journal of Experimental Psychology, 89*, 390–397.

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making, 27*, 696–713.

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.

Mellers, B. A., Ordóñez, L., & Birnbaum, M. H. (1992). A change-of-process theory for contextual effects and preference reversals in risky decision making. *Organizational Behavior and Human Decision Processes, 52*, 331–369.

Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review, 73*, 596–605.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38*, 332–382.

Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty, 6*, 235–254.

Tversky, A., & Fox, C. R. (1995). Weighing Risk and Uncertainty. *Psychological Review, 102*, 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., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review, 95*, 371–384.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *American Economic Review, 80*, 204–217.

Weber, B. J. (2007). The effect of losses and event splitting on the Allais paradoxes. *Judgment and Decision Making, 2*, 115–125.

Wu, G., Zhang, J., & Gonzalez, R. (2004). Decision under risk. In D. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 399–423). Oxford: Oxford University Press.