Behavioral Models of Decision Making under Risk

Michael H. Birnbaum

California State University, Fullerton

Mailing Address:

Dept. of Psychology
California State University H-830M
Box 6846
Fullerton, CA 92834-6846
USA

Email: mbirnbaum@fullerton.edu

Date: June 6, 2016

Behavioral Models of Decision Making under Risk

Michael H. Birnbaum

Some decisions are based on vague descriptions of the consequences of one's actions or on imprecise, uncertain, or ambiguous information concerning the probabilities of consequences contingent on one's alternative courses of actions. What job should I take? Whom should I marry? Should I undergo this operation to treat my medical condition? These decisions are made in the face of uncertainty. The term, *decision making under risk*, however, refers to situations in which a decision maker has valid information concerning the probabilities of consequences of the alternative courses of action. For example, should I buy a lottery ticket for $1 that has one chance in a million of paying $1 Million dollars? Researchers studying decision under risk are attracted to such decisions because gambles defined on events with known probabilities (such as tosses of fair coins or rolls of dice) allow one to manipulate clearly the important ingredients in the decision process itself, separated from the mechanisms by which beliefs about probability are formed.

Behavioral models of risky decision-making are theories that attempt to give empirically accurate descriptions of what people do when confronted with risky decision making problems. Whereas a normative model specifies what a person ought to do to stay consistent with certain principles of rationality, a behavioral model seeks to explain the empirical choices that people actually make, whether these choices are deemed rational or not.

In the simplest paradigm for study of decision making, we ask participants to make decisions among gambles stated in terms of probabilities to receive monetary consequences. For example, would you rather have $45 for sure, or would you prefer instead to play a risky gamble in which you have a 50-50 choice to win \$100 or \$0 based on the flip of a fair coin? Because the coin has a probability of ½ to be called correctly, you have a probability of ½ to win \$100 and a probability of ½ to win \$0. Most people prefer \$45 for sure to the risky gamble. This systematic preference for the sure thing contradicts a rule called expected value, which was once thought to be a rational principle a person should follow.

Expected Value

The expected value (EV) of a gamble is the mean value of the outcomes, weighted by their probabilities. Suppose a random process has n possible mutually exclusive and exhaustive outcomes, and let gamble $G = (x_1, p_1; x_2, p_2; x_3, p_3; \ldots; x_i, p_i; \ldots; x_n, p_n)$ represent a gamble with probability p_i to receive consequence x_i, where x_i is the monetary consequence if outcome i occurs. Because the outcomes are mutually exclusive and exhaustive, $\Sigma\, p_i = 1$. We can define the EV of gamble G as follows:

$$\mathrm{EV}(G) = \Sigma\, p_i\, x_i$$

The gamble based on a coin flip is denoted G = (\$100, ½; \$0, ½), and G has an EV of \$50 = (.5)(\$100) + (.5)(\$0). So, on average, a person who could play G infinitely many times would win \$50, but on each trial, the person would win either \$0 or \$100.

Risk Aversion and St. Petersburg Paradox

EV seemed a reasonable objective measure to scholars in the 18th century, so they considered it paradoxical that when given a choice, people did not always prefer the option with the higher EV. For example, many people prefer \$45 for sure to G = (\$100, ½; \$0, ½), , even though the sure thing has a lower EV ($45) than the gamble ($50). When people prefer a sure thing to a gamble with the same or higher EV, they are said to be "risk averse."

Risk aversion seemed puzzling, especially when scholars realized that they could construct gambles with infinite value, and people preferred even small amounts of cash to such gambles. For example, suppose we toss a coin and if it is Heads, you win $2, but if it is Tails, we toss again. This time, if it is Heads, the payoff is $4 and if Tails, we toss again. Each time that tails occurs, the prize for Heads on the next toss doubles. The expected value of this gamble is as follows:

$$EV = \$2(1/2) + \$4(1/4) + \$8(1/8) + \$16(1/16) + \ldots = \infty$$

If a person conformed to EV, she should prefer this gamble to any finite amount of money one might offer—yes, a person should prefer playing this gamble once to all of the money in Switzerland. Yet, most people say they would prefer $20 for sure to one chance to play this gamble, even though the gamble has infinite expected value.

This preference for the sure thing over such gambles (with infinite EV) is now called the *St. Petersburg paradox*, which was discussed in a classic paper by Bernoulli (1738), who presented his paper in St. Petersburg. Bernoulli said it was not necessarily rational to follow EV, but instead to choose the option with the best expected utility.

Expected Utility

Bernoulli (1738/1954) provided an explanation of risk aversion that solved the original versions of the St. Petersburg paradox. This theory proposed that the utility of money is not necessarily equal to its objective value, but might, instead, be a negatively accelerated function of money. Let $u(x)$ represent the utility, or "moral" value, of a certain amount of cash, x. Define Expected Utility (EU) as follows:

$$\text{EU}(G) = \Sigma\, p_i\, u(x_i) \qquad (1)$$

Where $u(x_i)$ is the utility of objective cash value x_i. Expected utility theory is the theory that people prefer A over B if and only if EU(A) > EU(B).

Bernoulli theorized that utility of money might be a logarithmic function of money, but he acknowledged that a power function, such as the square root function, might also work.

EU theory could explain not only risk aversion and the original St. Petersburg paradox, but it could also explain why a pauper who was given a lottery ticket should be happy to sell it for less than EV, and why a rich person should be happy to buy it at the same price. For example, suppose utility is given by the following:

$$u(x) = x^{0.5}$$

Imagine a pauper whose total wealth is just \$50, who is given a choice between S = \$45 for sure, and G = 50-50 gamble to win \$100 or \$0. According to EU theory, if $u(x) = x^{0.5}$, the utility of choosing the sure thing, S is u(\$50+\$45) = u(\$95) = 9.75. The EU of choosing gamble G is u(\$50 + \$100)(0.5) + u(\$50) = 9.66. Because EU(S) > EU(G), the theory says the pauper would prefer \$45 for sure over gamble G. The pauper who was given a lottery ticket (a chance to play gamble G) would be happy to sell it for \$45.

Now consider a richer person whose total wealth is \$1000, who is deciding whether to buy gamble G from the pauper. The utility of Q, the status quo (to not buy) is u(\$1000) = 31.62. The utility of B, the option to buy the gamble for \$45 from the pauper has expected utility of u(\$1000 + \$100 - \$45)(0.5) + u(\$1000 - \$55)(.5) = 31.69. Because EU(B) > EU(Q), this person should prefer to buy the gamble for \$45. This example shows that even if both people have the same utility function (but different levels of wealth), they can both improve their individual utilities by trading.

It is also possible that some people have different utility functions from others, reflecting different attitudes toward risk. For example, if a venturesome person had $u(x) = x^2$, then that person would prefer the risky gamble to a sure thing with the same EV, and would be called "risk-seeking." Such a risk-seeking person would out-bid the wealthy but risk averse person to buy gamble *G*, and would even be willing to buy this gamble at a price exceeding $50.

Von Neumann and Morgenstern (1947) showed that expected utility theory could be deduced from four basic axioms of preference and proved that if these axioms held, utility could be measured on an interval scale. This theory was taken as the definition of what a rational person should do when confronted with decisions under risk.

Much of economic theory had been deduced from the theory that people are rational but may differ in their utilities or tastes. For a time, it was also thought that people are rational, which led to the view that classic economic theory not only described what a rational economic actor would do, but was also descriptive of actual behavior of individuals. EU theory was treated as if it was not only rational but also descriptive of human economic behavior. However, both the assumption of rationality of EU and the assumption that people are rational came into question when Allais proposed his paradoxes.

Allais Paradoxes

In the early 1950s, Allais criticized EU theory from both descriptive and normative perspectives. He developed paradoxes that have generated continued discussion in the scientific literature that continue to this day. Consider the following two choice problems:

Problem 1:

A: ($1 Million, .11; $0, .89)

B: ($2 Million, .10; $0, .90)

Problem 2:

C: ($1 Million, with certainty)

D: ($2 Million, 0,1; $1 Million, .89; $0, .01)

According to EU theory, a person should prefer *C* over *D* if and only if (iff) she prefers *A* over *B*; however, many people prefer *C* over *D* and *B* over *A*, contrary to the theory. This paradox is known as the "constant consequence" paradox because .89 probability to win $0 is common to both *A* and *B*, which has been changed to a common consequence of .89 to win $1 Million in *C* and *D*.

A "constant ratio" paradox was also developed, which can be illustrated by the following choices:

Problem 3:

E: $3000 for sure

F: ($4000, 0.8; $0, 0.2)

Problem 4:

G: ($3000, .25; $0, .75)

H: ($4000, .20; $0, .80)

According to EU theory, a person should prefer *E* to *F* if and only if she prefers *G* to *H*; however, many people prefer *E* to *F* and prefer *H* to *G*. The constant ratio refers to the fact that the probabilities to win in *G* and *H* of Problem 4 are one fourth of those in *E* and *F* of Problem 3. These paradoxes refuted EU as a descriptive model of how people choose between risky gambles. To Allais (1979), these paradoxes reflected shortcomings of EU as a rational model as well.

Subjectively Weighted Utility and Prospect Theory

Ward Edwards (1953) used a subjectively weighted model to account for the Allais paradoxes. According to the model of Edwards, the value of a gamble is given by the following:

$$PV(G) = \Sigma\, w(p_i)\underline{u}(x_i) \qquad (2)$$

Where PV(*G*) is the prospect value of a gamble and $w(p_i)$ is the weight of the probability. Amos Tversky, a former student of Edwards, and Kahneman published a variant of this model in *Econometrica* under the name "prospect theory" (Kahneman & Tversky, 1979; Kahneman, 2003). Equation 2 is sometimes called "stripped" prospect theory.

This formulation could account for the Allais paradoxes, but it made some strange predictions that seemed unrealistic. For example, it predicted that people should prefer gamble *I* = ($100, .01; $100, .01; $99, .98) to *J* = ($102; .5; $101, .5), even though every outcome of *J* is better than any outcome of *I*. Because it seemed unlikely that people would violate stochastic dominance (e.g., choose *I*) in such cases, Kahneman and Tversky postulated editing rules that people supposedly used to avoid such implications of this model, and they postulated other restrictions and exceptions to Equation 2.

Rank dependent weighting was proposed (Quiggin, 1985) as a way to account for the Allais paradoxes without violating stochastic dominance. Luce and Fishburn (1991) axiomatized rank-and sign-dependent utility (RSDU). Tversky and Kahneman (1992) adopted a version of a rank- and sign-dependent model and called it cumulative prospect theory (CPT). According to RSDU or CPT, the value of a gamble on strictly non-negative consequences is given by the following:

$$\text{CPV}(G) = \sum [W(P_i) - W(Q_i)]\underline{u}(x_i) \tag{3}$$

Where W is a strictly monotonic function from $W(0) = 0$ to $W(1) = 1$ that assigns decumulative weight to decumulative probability, P_i is the decumulative probability to win x_i or more and Q_i is the probability to win strictly more than x_i.

This model always satisfies stochastic dominance and it also satisfied other principles that had required editing rules in original prospect theory. CPT could account for the Allais paradoxes by means of an inverse-S shaped decumulative weighting function. This function assigned more weight to branches leading to smallest and largest consequences than to branches leading to intermediate ones. For a time, CPT appeared a better description than expected utility theory, but it had not been tested against another approach that had been proposed in the 1970s that shared some features of rank-dependent weighting, but differed in some important ways.

Configural Weighting Models

Birnbaum (1974; Birnbaum & Stegner, 1979) proposed configural weight models in which the rank of a stimulus affects its weight. Those aspects of a stimulus that are more unfavorable often received greater weight. Although these models had much in common with the models later introduced as "rank dependent utility", the configural weight models do not always satisfy stochastic dominance.

Configural weighting provides a different interpretation of risk aversion than found in EU theory: according to EU theory, risk aversion is produced by curvature of the utility function; according to configural weighting theory, however, risk aversion or risk seeking is mainly produced by over or under weighting of the lower valued consequences or aspects of a gamble or stimulus.

Consider the simple case of a 50-50 gamble to win either x or y, where $x > y \geq 0$. The *transfer of attention exchange* (TAX) model for this gamble can be written as follows:

$$\text{TAX}(G) = (.5 + \omega)u(x) + (.5 - \omega)u(y)$$

Where ω is the configural weight transferred from the lower-valued to the higher valued consequence or aspect of the gamble or stimulus ($-0.5 \leq \omega \leq 0.5$). If $\omega = 0$, TAX reduces to EU; if $\omega = 0.5$, it becomes a maximum model, and with $\omega = -0.5$, it becomes a minimum model. For gambles on small amounts of cash ($x <$ \$150), with college students, one can approximate $u(x)$ as $u(x) = x$, and $\omega = -1/6$. With those parameters, a person would prefer \$40 for sure to the 50-50 gamble to win \$100, and would prefer the gamble to \$20 for sure, being indifferent between the gamble and \$33 for sure.

For binary gambles of the form, $G = (x, p; y)$, $x > y \geq 0$, the TAX model can be written as follows:

$$\text{TAX}(G) = [au(x) + bu(y)]/(a + b)$$

Where a and b are the weights of the higher and lower consequences, which have utilities of $u(x)$ and $u(y)$, respectively. The weights are given as follows:

$$a = t(p) - \delta t(p)/3;$$

$b = t(1 - p) + \delta t(p)/3;$

where $t(p)$ is a function of p, usually approximated as a power function, and $\delta > 0$ is a constant reflecting the transfer of weight (attention) from the higher valued consequence to the lower valued consequence. When the transfer goes the other direction, $\delta < 0$, and one replaces $t(p)$ with $t(1 - p)$ in the above equations.

With three-branch gambles of the form, $G = (x, p; y, q; z, 1 - p - q)$, $x > y > z \geq 0$, the model is still a weighted average, $\text{TAX}(G) = [Au(x) + Bu(y) + Cu(z)]/(A + B + C)$, where the weights (for branches with highest, middle, and lowest consequences) are as follows for a person who places greater weight on lower valued consequences:

$A = t(p) - 2\delta t(p)/4$

$B = t(q) + \delta t(p)/4 - \delta t(q)/4$

$C = t(1 - p - q) + \delta t(p)/4 + \delta t(q)/4$

Previous research has shown that modal choices by undergraduates for gambles involving small positive values can be roughly approximated by with $t(p) = p^{0.7}$, $u(x) = x$, and $\delta = 1$. Although these "prior" parameters have done fairly well in predicting group data for the last twenty years, data fitting shows that the estimated utility function should be negatively accelerated, especially when consequences cover a large range of values or include large values.

There are two aspects of the weights that deserve emphasis: First, the transfer of weights has the implication that risk aversion or risk seeking can be explained by greater or reduced weight on the lower valued consequence rather than via the utility function.

Second, the weighting of branches need not satisfy coalescing, which is the assumption that splitting a branch of a gamble would not affect its utility. For example, coalescing implies that *A* = ($96, .85; $96, .05; $12) should have the same utility as *B* = ($98, 0.9; $12). Note that *A* and *B* are (objectively) the same; *B* is called the *coalesced* form of the gamble, and *A* is one of many possible *split* forms of the same gamble. Instead, splitting a branch increases the weight given to the consequences of the split branch. This implication follows from the fact that $t(p)$ is negatively accelerated, like many other psychophysical functions. In this averaging model, splitting the branch leading to the highest consequences tends to make a gamble better (subjectively) and splitting the branch leading to the lowest consequence tends to make a gamble seem worse.

Differences in the properties and predictions between the configural weight models and RSDU models including CPT were identified and tested by Birnbaum in a series of experiments that refuted this class of models as descriptive of decision making (Birnbaum, 2004a; 2004b; 2006). The configural weight models, fit to previous data, correctly predicted where to find new violations, which Birnbaum (2008b) called "new paradoxes" because these critical properties refuted CPT in the same way that Allais paradoxes refuted EU; that is, they lead to contradictions in the model that cannot be explained by revising parameters or functions in the model. Two of these critical tests among these models are reviewed in the next sections (others are reviewed in Birnbaum, 2008b, 2008c).

Violations of Stochastic Dominance

If the probability to win a prize of *x* or greater in gamble *F* is always at least as high and sometimes higher than the corresponding probability in gamble *G*, we say that gamble *F* dominates gamble *G* by first order stochastic dominance. According to rank-and sign-dependent utility theories, including CPT and EU, first order stochastic dominance must be satisfied. The configural weight models, however, imply that special choice problems can be constructed in which people will violate stochastic dominance.

Birnbaum and Navarrete (1998) tested choice problems such as the following that were predicted by the TAX and RAM models to violate stochastic dominance:

Problem 5:

K: ($96; .90; $14, .05; $12, .05)

L: ($96, .85; $90, .05; $12, .10)

Birnbaum and Navarrete (1998) found that about 70% of undergraduates choose *L* over *K*, even though *K* dominates *L*. Note that the probability to win $96 or more is higher in *K* than *L*, the probability to win $90 or more is the same, the probability to win $14 or more is higher in *K* than *L,* and the probability to win $12 or more is the same. There have been dozens of studies reporting similar, substantial violations of stochastic dominance in choice problems of this type, using different types of participants, different types of cash incentives, different types of probability mechanisms, different formats for presenting choice problems, and different types of event framing (Birnbaum, 2004a; 2004b; 2006; 2007; 2008b; Birnbaum & Bahra, 2012). These violations show that no form of rank-and sign-dependent utility function, including CPT, can be considered as a descriptive model of risky decision making, but they were predicted by the configural weight models that were used to design the experiment.

Dissection of the Allais Paradox

Birnbaum (2004) noted that constant consequence paradigm of Allais can be decomposed into three simpler properties: transitivity, coalescing, and restricted branch independence. *Transitivity* is the assumption that one prefers *A* to *B* and prefers *B* to *C*, then one should prefer *A* to *C*. *Coalescing* is the assumption that if two branches of a gamble lead to the same consequence, they can be combined by adding the probabilities, without changing utility. For example, in Problem 1, the gamble, *A* = ($1 Million, .11; $0, .89) is assumed to be identical in utility to A_S = ($1 Million, .10; $1 Million, .01; $0, .89), where A_S is one of the "split" forms of *A*, the same gamble in "coalesced" form.

Restricted branch independence is the assumption that if two gambles with the same number of branches and same probability distribution over those branches have a common consequence on a branch, the common consequence can be changed without altering the preference. For example, A_S = ($1 Million, .10; $1 Million, .01; $0, .89) is preferred to B_S = ($2 Million, .1; $, .01; $0, .89) if and only if C_S = ($1 Million, .10; $1 Million, .01; $1 Million, .89) is preferred to D_S = ($2 Million, .10; $0, .01; $1 Million, .89) because the common branch of .89 to win $0 has been changed to a common branch of .89 to win $1 Million.

If a person satisfied transitivity, coalescing, and restricted branch independence (all implied by EU), that person would not display the constant consequence paradox of Allais (Birnbaum, 2004).

Consider the choice problems in Table 1. According to EU theory, the preference should be the "same" in all six choice problems, in the sense that *A* preferred to *B* if and only if (iff) A_S preferred to B_S, iff C_S preferred to D_S, iff *C* preferred to *D*, iff E_S preferred to F_S, and iff *E* preferred to *F*.

Insert Table 1 about here.

Table 1. Six choice problems dissecting Allais paradox into tests of coalescing and restricted branch independence. According to coalescing, Choices No. 1 and 1s, 2 and 2s, and 3 and 3s are equivalent choice problems. According to restricted branch independence, Choices No. 1s, 2s, and 3s should all be either "Safe" or they should all be "Risky", but one should not switch systematically.

No.	"Safe"	"Risky"
1	*A*: ($1M, .11; $0, .89)	*B*: ($2M, .10; $0, .89)
1s	A_s: ($1M, .10; $1M, .01; $0, .89)	B_s: ($2M, .10; $0, .01; $0, .89)
2s	C_s: ($1M, .10; $1M, .01; $1M, .89)	D_s: ($2M, .10; $0, .01; $1M, .89)
2	*C*: $1M for sure	*D*: ($2M, .10; $1M, .89; $0, .01)
3s	E_s: ($1M, .10; $1M, .01; $2M, .89)	F_s: ($2M, .10; $0, .01; $2M, .89)
3	*E*: ($2M, .89; $1M, .11)	*F*: ($2M, .99; $0, .01)

Original prospect theory, CPT, and TAX do not agree with EU, and they do not agree with each other, so one can compare all three theories by testing this "dissection" of the Allais paradox (Birnbaum, 2004; 2007). The predictions of the theories are shown in Table 2. Original prospect theory implies restricted branch independence and attributes the paradox to violations of coalescing. That is, OPT implies A_S is preferred to B_S, iff C_S is preferred to D_S, iff E_S is preferred to F_S. To explain an Allais paradox such as a reversal between the choice between *A* and *B* and between *E* and *F*, there must be a reversal either between choices of *A* versus *B* and A_S versus B_S, or between the choices of E_S versus F_S, and *E* versus *F*. OPT also had editing rules of combination and cancellation that imply coalescing and restricted branch independence, respectively, so OPT could mimic EU by invoking these editing rules, in which case the model would not show the Allais paradoxes.

Insert Table 2 about here.

Table 2. Comparison of Decision Theories.

	Branch Independence	
Coalescing	Satisfied	Violated
Satisfied	EUT (OPT*/CPT*)	CPT
Violated	OPT	CWT

Notes: EUT = Expected Utility Theory; CPT = Cumulative Prospect Theory; OPT= Original Prospect Theory; CWT = Configural Weight Theory (TAX). *The editing rules of combination and cancellation produce satisfaction of coalescing and restricted branch independence, respectively.

In contrast, CPT assumes coalescing and attributes the Allais paradox to violations of restricted branch independence; thus, *A* is preferred to *B* iff A_S is preferred to B_S,; and E_S is preferred to F_S, iff *E* is preferred to *F*. If cancellation was invoked, CPT could also mimic EU.

Configural weight models such as TAX violate coalescing and they imply opposite violations of restricted branch independence from those required by CPT to account for the Allais paradoxes. According to this model, it should be possible to construct choice problems in which the Allais paradox would be reversed if the choices were presented in canonical split form. That is, when the probabilities on ranked branches are equal and the number of branches is minimal, as in choices A_S versus B_S, and E_S versus F_S,.

Empirically, there are strong violations of both coalescing and of restricted branch independence, and the violations of restricted branch independence are opposite the direction required by CPT to account for the Allais paradox (Birnbaum, 2004, 2007). Thus, EU and both versions of prospect theory can be rejected.

A number of studies have now been completed testing between configural weight models and CPT investigating these and other critical behavioral properties that can be used to distinguish between these models. The results strongly refute both versions of prospect theory in favor of the predictions made by the configural model (Birnbaum, 2008b; Birnbaum & Bahra, 2012a).

Priority Heuristic

Brandstätter, Gigerenzer, & Hertwig (2006) devised a heuristic model that emulates CPT. This model was based on the lexicographic semiorder employed by Tversky (1966) to describe intransitive preferences that Tversky reported for a small number of selected individuals. According to the priority heuristic (PH), a person first compares lowest consequences of a gamble and chooses the gamble with the higher lowest consequence if they differ by more than 10% of the largest consequence in either gamble, rounded to the nearest prominent number. When the lowest consequences are not sufficiently different, the person chooses 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 differed by less than 0.1, the person is theorized to next compare the highest prizes and choose by that criterion, if they differ sufficiently. When there are more than two branches and the first three comparisons yield no decision, the person next compares the probabilities to win the highest prize and decides on that basis alone, if there is any difference. And if all four criteria yield no decision, the person chooses randomly, without examining anything else. At each stage, the decision is based on only one reason, which is the contrast on one dimension.

A claim was made that the PH fit published choice data as well or better than EU, CPT, or TAX, but this claim was challenged and shown to hold only with selected data and only when certain assumptions are forced onto theories that do not make those assumptions (Birnbaum, 2008a). The PH model was constructed to account for the Allais paradoxes in original form, but it could not account for the dissection of the Allais paradoxes, nor for violations of stochastic dominance, nor for violations of restricted branch independence.

The PH implies systematic violations of transitivity. For example, if A = ($5.00, .29; $0), C = ($4.50, .38; $0), and E = ($4.00, .46; $0), then the PH predicts that the majority should prefer A to C, and prefer C to E, and yet prefer E to A. But studies by Birnbaum & Gutierrez (2007), Regenwetter, Dana, and Davis-Stober (2011), and Birnbaum and Bahra (2012b), among others, found that majority preferences did not show the predicted choices by the PH. In fact, the PH predicted only 30% of the modal choices correctly in Birnbaum and Gutierrez (a random coin toss would correctly predict 50%). This model also does significantly worse than chance in predicting violations of restricted branch independence, because it predicts the opposite pattern of violations from what is observed (Birnbaum & Bahra, 2012a).

The PH implies that attributes or dimensions of a stimulus do not combine, nor do they interact, but experimental tests of combination and interaction showed evidence that people integrate information between dimensions and that the dimensions interact. For example, consider the following two choice problems:

Problem 6:

X = ($100, 0.9; $5, 0.1)

Y = ($50, 0.9; $20, 0.1)

Problem 7:

X' = ($100, 0.1; $5, 0.9)

Y' = ($50, 0.1; $20, 0.9)

According to the PH, a person should choose *Y* and *Y'* because the probabilities are the same and the lowest consequences are better by the same amount in both gambles. According to another lexicographic semiorder, a person might choose *X* and *X'*, if they examined the highest consequences first. Because the probabilities are the same in both gambles within each choice, probability should not make any difference in these models. However, most people choose *X* over *Y* in Problem 6 and choose *Y'* over *X'* in Problem 7, contrary to any lexicographic semiorder model. These violations also contradict similarity models that decide by comparing contrasts between components but do not postulate that components interact (Birnbaum, 2008c, 2010).

The perceived relative arguments model (Loomes, 2010), like the priority heuristic and regret theory (Loomes & Sugden, 1982; Loomes, Starmer, & Sugden, 1991), can violate transitivity. This model assumes that people make choices by combining contrasts between the components. However, empirical studies of predicted intransitivity by regret theory and perceived relative arguments model have not confirmed its predictions, and the model also fails to account for violations of restricted branch independence (Birnbaum & Diecidue, 2015).

In principle, violations of transitivity, if substantial and systematic, would rule out the class of models that includes EV, EU, CPT, and TAX. Therefore, it would be extremely important to know if stimuli can be found that produce reliable violations of transitivity. I am not aware of any experiment in which a majority of participants has been shown to exhibit reliable intransitive cycles.

Decisions from Description and from Experience

Hertwig, Barron, Weber, & Erev (2004) contrast two paradigms for decision making problems. The first method asks people to make a single decision based on descriptions of the relevant probabilities and consequences, and the other method involves learning of probabilities based on experience with a sequence of events representing some stochastic process. With description, many people will prefer a small chance at a large prize to a sure thing with the same expected value. For example, many people prefer M = ($100, .01; $0) over N = $1 for sure, based on a description.

Such risk-seeking behavior for small probabilities to win positive consequences is consistent with OPT, CPT, and TAX, given suitable parameters. However, when people are asked to sample from the two options, and then asked to make a choice, they often choose the safe option over the risky gamble. Hertwig, et al. (2004) argued that perhaps different theories of decision making might be required for these two types of situation. They note that learning and perception of probabilities might be overly influenced by the most recent events in a sequence.

However, Fox and Hadar (2006) noted that from the perspective of experience, many people who drew small samples might experience M as $0 always occurs (since the unlikely event of $100 might never occur in a small sample), and N as a always paying $1; they never experience the population, so subjectively, the choice was between always nothing versus always $1. In many of the studies done in this field, sampling is left to the participant and to chance, so the experience has not been constrained to match the description.

Glöckner, Hilbig, Henniger, and Fiedler (2016) present a current review of the literature on description versus experience, a reanalysis of earlier studies, and new experiments designed to disentangle different interpretations. They conclude that sampling and regression effects are important components of the previous studies, but they argue that other factors (such as uncertainty) play roles as well. For example, how does one learn from a brief experience that something is a "sure thing?" When one hears, "you win $50 no matter what color you draw from the urn," it denotes a sure thing. This case is different from "you win $50 only if you draw a red ball from the urn", and then after 15 trials a person has experienced only red balls. There is still the chance that other colors may be in the urn that have not yet been sampled.

One factor that has not yet been addressed in this literature on experience versus description that has been considered by some in the description literature is the role of error or variability of response in producing choice behavior.

Models of Error or Variability

When a person is presented the same choice problem on two occasions, the same person will often make a different choice response on the two trials. For example, consider the next two choice problems:

Problem 8:

R = ($98, .10; $2, .90),

S = ($40, .20; $2, .80);

Problem 9:

R' = ($98, .90; $2, .10);

S' = ($98, .80; $40, .20).

Problems 8 and 9 were included, separated by a number of intervening trials, among a list of 31 choice problems. Following a brief intervening task of about 10 minutes, the same people were asked to respond to the same choice problems a second time. It was found that the same people made different responses 20% of the time on Problem 9, and 31% reversed preferences on two presentations of Problem 8. According to EU, a person should prefer *R* over *S* if and only if she prefers *R'* over *S'*. But if the same person can change responses when Problem 8 is presented twice, should we be surprised if that same person made different responses on Problem 8 and 9?

In the past, researchers argued that if significantly more people chose *R* in Problem 8 and *S'* in Problem 9 than the number who made the opposite pattern of reversal (*S* and *R'*), then the "significant" difference meant one should reject EU. However, it has recently been shown that if different choice problems have different rates of error, then such asymmetry of reversals could occur even if EU held true. The idea that inherent variability or errors in choice might produce some or all of the apparent violations of behavioral properties such as transitivity has been an important focus of recent research (Regenwetter, Dana, & Davis-Stober, 2011; Birnbaum, 2013; Birnbaum & Bahra, 2012a; Carbone & Hey, 2000; Loomes, 2005; Wilcox, 2008).

A family of models known as "true and error" models has been developed, based on the idea that one can estimate the error component from preference reversals by the same person to the same choice problems within a brief session. These models allow that a person's "true" preferences may have variability between sessions (blocks of trials), due to such factors as changing parameters or changing models, and they allow separation of such variability due to a mixture of models from variability produced by "error" that produces reversals within a session. They allow each choice problem to have a different rate of error, and they allow different people to have differing amounts of noise or unreliability in their responses.

When these models have been applied to repeated judgments, it has been found that the violations of EU, as in the Allais paradoxes, violations of CPT, as in the "new paradoxes", and violations of the priority heuristic, as in the tests of interactive independence, cannot be attributed to error (Birnbaum, 2008b; 2008c; 2010). On the other hand, violations of transitivity have been found to be of low frequency when the inherent variability of the data are fit by the true and error model.

Concluding Comments

The field of risky decision-making is one of the oldest topics in behavioral science and has influenced both psychology and economics. Over the years, new models and new evidence has accumulated to refute some theories in favor of others. When the new evidence violated a currently popular model, the findings were called "paradoxes" or "anomalies". Data have shown that EV, EU, SWU, OPT, and CPT can be rejected based on violations of critical properties. Intransitive models such as regret theory, lexicographic semiorders, and the priority heuristic, have not yet been able to show where to find the predicted intransitive preference cycles, nor have they been successful in predicting results of new experiments designed to test them. Configural weight models, such as TAX, remain consistent with the major phenomena. As new theories are developed, new tests are designed and new information is gained that constrains the possible theories that can represent this rich field of human behavior.

References

Allais, M. (1953). Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine, *Econometrica, 21*, 503-546.

Allais, M. (1979). The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school (1952). In M. Allais & O. Hagen (Eds.), *Expected Utility Hypotheses and the Allais Paradox*. Dordrecht, Netherlands: Reidel/Springer. 27-145.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*, 23-36. (English translation of 1738 paper).

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. (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,* 253-260.

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. (2010). Testing lexicographic semiorders as models of decision making: Priority dominance, integration, interaction, and transitivity. *Journal of Mathematical Psychology, 54*, 363-386.

Birnbaum, M. H. (2013). True-and-error models violate independence and yet they are testable. *Judgment and Decision Making, 8,* 717-737.

Birnbaum, M. H., & Bahra, J. P. (2012a). Separating response variability from structural inconsistency to test models of risky decision making, *Judgment and Decision Making, 7*, 402-426.

Birnbaum, M. H., & Bahra, J. P. (2012b). Testing transitivity of preferences in individuals using linked designs. *Judgment and Decision Making, 7*, 524-567.

Birnbaum, M. H., & Diecidue, E. (2015). Testing a class of models that includes majority rule and regret theories: Transitivity, recycling, and restricted branch independence. *Decision, 2,* 145-190.

Birnbaum, M. H., & Gutierrez, R. J. (2007). Testing for intransitivity of preferences predicted by a lexicographic semiorder. *Organizational Behavior and Human Decision Processes, 104*, 97-112.

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

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

Carbone, E., & Hey, J. D. (2000). Which error story is best? *Journal of Risk and Uncertainty, 20*, 161-176.

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.

Glöckner, A., Hilbig, B., Henninger, F., & Fiedler, S. (2016). The reversed description-experience gap: Disentangling sources of presentation format effects in risky choice. *Journal of Experimental Psychology: General, 145,* 486-508.

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

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.

Loomes, G. (2005). Modeling the stochastic component of behavior in experiments: Some issues for interpretation of data. *Experimental Economics, 8*, 301-323.

Loomes, G. (2010). Modeling choice and valuation in decision experiments. *Psychological Review, 117*, 902-924.

Loomes, G. & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal, 92*, 805-824.

Loomes, G., Starmer, C., & Sugden, R. (1991). Observing violations of transitivity by experimental methods. *Econometrica, 59*, 425-440.

Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches.* Mahwah, NJ: Lawrence Erlbaum Associates.

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.

Marley, A. A. J., & Luce, R. D. (2005). Independence properties vis-à-vis several utility representations. *Theory and Decision, 58*, 77-143.

Quiggin, J. (1985). Subjective utility, anticipated utility, and the Allais paradox. *Organizational Behavior and Human Decision Processes, 35*, 94–101.

Quiggin, J. (1993). *Generalized expected utility theory: The rank dependent model*. Boston: Kluwer.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of preferences. *Psychological Review*, *118*, 42-56.

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

Von Neumann, J. & Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press. (2nd Edition).

https://archive.org/details/theoryofgamesand030098mbp

Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox and G. W. Harrison (eds), *Research in Experimental Economics Vol. 12: Risk Aversion in Experiments* (pp. 197-292). Bingley, UK: Emerald.