Journal of Mathematical Psychology 75 (2016) 183–193

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Empirical evaluation of four models of buying and selling prices of gambles

Michael H. Birnbaum[a,*], Sherry Yeary[a], R. Duncan Luce[b], Li Zhao[b]

[a] California State University, Fullerton and Decision Research Center, United States
[b] University of California, Irvine, United States

ARTICLE INFO

Article history:
Available online 14 June 2016

Keywords:
Buying price
Choice
Configural weighting
Contingent valuation
Cumulative prospect theory
Endowment effect
Utility
Loss aversion
Preference reversals
Selling price
Viewpoint
Willingness to accept (WTA)
Willingness to pay (WTP)

ABSTRACT

Judges assigned values to gambles from viewpoints of buyers (willingness to pay) and sellers (willingness to accept). Consistent with previous results, selling prices exceed buying prices, and these two judgments are not monotonically related to each other. There are systematic violations of consequence monotonicity when the consequence of zero is increased to a small positive value. Models based on loss aversion combined with cumulative prospect theory (CPT) do not give accurate accounts of the data. In particular, judgments violate complementary symmetry, which is implied by third-generation prospect theory. In addition, there are violations of first order stochastic dominance in judgments of three-branch gambles. Models based on the theory of joint receipts by R. D. Luce fit better than third-generation prospect theory, but the best-fitting of six such models does not give an adequate account of judgments involving a lowest consequence that might be zero or positive. Two configural weight models give better fits to the data using the same number or fewer parameters estimated from the data.

When people are asked to judge the highest price they would pay to buy something, they report a much lower value than when they are asked to judge the least they would accept to sell the same thing. This finding has been observed for both goods of uncertain or ambiguous value, such as used cars, and for gambles with defined probabilities and consequences (Birnbaum & Stegner, 1979; Coombs, Bezembinder, & Goode, 1967). The discrepancy between judged willingness to pay (WTP) and willingness to accept (WTA) is surprisingly large from the perspective of classical expected utility theory in economics (Horowitz & McConnell, 2002; Knetsch & Sinden, 1984).

The same phenomenon has been discussed in different isolated segments of the scientific literature and attributed to different sources. Birnbaum and Stegner (1979) referred to it as an effect of the judge's point of view and tested rank-affected configural weight models. Thaler (1980) used the term "endowment effect" and theorized that it might be due to loss aversion. It is also known as a contingent valuation effect (Irwin, Slovic, Lichtenstein, & McClelland, 1993). Luce (1991, 2000) theorized that the difference could be described by a joint receipt model in which buying involved the joint receipt of a positive gain of the item purchased and the loss of the price paid and selling involved a loss of the item sold and a gain of the purchase price.

This article brings together and compares models that have arisen in different branches of the scientific literature. We report an experiment that evaluates them as empirical representations. Before presenting the models, it is useful to cite briefly three approaches to this topic, because recent reviews have focused on only part of the literature. For example, Erickson, Marzilli, and Fuster (2014), in the *Annual Review of Economics*, do not cite research on configural weighting and the point of view effect, and although Morewedge and Giblin (2015) take a broader perspective, like Erickson et al., they do not cite Duncan Luce, despite the relevance of his work.

The three groups of models considered here are (1) prospect theory loss aversion models (Birnbaum & Zimmermann, 1998; Kahneman, Knetsch, & Thaler, 1990; Schmidt, Starmer, & Sugden, 2008; Tversky & Kahneman, 1991, 1992) which assume that buyers and sellers experience different patterns of losses and gains, even when buying or selling objects or lotteries that are strictly positive; (2) joint receipt models (Luce, 1991, 2000), which assume that the price paid to buy a lottery or to sell one are integrated into

* Correspondence to: Department of Psychology, H-830M, California State University, Fullerton, P.O. Box 6846, Fullerton, CA 92834-6846, United States. Tel.: +1 657 278 2102; fax: +1 657 278 7134.
E-mail address: mbirnbaum@fullerton.edu (M.H. Birnbaum).

http://dx.doi.org/10.1016/j.jmp.2016.05.007
0022-2496/

the consequences by means of a joint receipt operation; and (3) configural weight models (Birnbaum, 1982; Birnbaum, Coffey, Mellers, & Weiss, 1992; Birnbaum & Stegner, 1979), which assume that buyers assign higher weight to lower-valued consequences, estimates, or attributes of an entity than do sellers.

1. Theoretical approaches

1.1. Cumulative prospect theory and loss aversion

Let $g = (x, p; y)$ represent a binary gamble (lottery) with probability p to receive x and otherwise receive y, where $x > y$. The CPT model represents the utility of g as follows:

$$U(x, p; y) = \begin{cases} U(x)W^{+}(p) + U(y)[1 - W^{+}(p)] & \text{if } (x \geq y \geq 0) \\ U(x)W^{+}(p) + U(y)W^{-}(1 - p) & \text{if } (x \geq 0 \geq y) \\ U(x)[1 - W^{-}(1 - p)] + U(y)W^{-}(1 - p) & \text{if } (0 \geq x \geq y). \end{cases} \quad (1)$$

In these expressions, $U(x)$ is the utility (value) function; we assume $U(0) = 0$; $W^{+}(p)$ and $W^{-}(p)$ are the weighting functions for a gain or loss received with probability p, respectively; $W^{+}(0) = W^{-}(0) = 0$ and $W^{+}(1) = W^{-}(1) = 1$.

Loss aversion has most commonly been modeled as follows:

$$U(-x) = -\lambda U(x), \quad x > 0, \quad (2)$$

where $U(-x)$ is the utility of a loss of x ($x > 0$), and λ is the factor by which losses are said to "loom larger" than gains, when $\lambda > 1$.

The term, *loss aversion*, unfortunately, has been used in two different ways that may have created confusion in the literature: first, it has been used to refer to the behavioral phenomenon of risk aversion for certain mixed gambles; second, it has been used to refer to the theory that empirical findings of risk aversion for such mixed gambles are due to the shape of the utility function. A rival explanation for the behavioral phenomena is that negative consequences receive greater weight (e.g., Birnbaum & Bahra, 2007), rather than more extreme utility.

It has been reported that choices satisfy *reflection*, which is the behavioral property that if gamble $g = (x, p; y)$, where $x > y > 0$, is preferred to gamble f, then gamble $-f$ is preferred to gamble $-g$, where $-g = (-x, p; -y)$. If Eq. (2) holds and if $W^{+}(p) = W^{-}(p)$, then reflection would be satisfied (Tversky & Kahneman, 1992). A more general form of "loss aversion" that need not satisfy reflection is considered in the discussion.

1.1.1. Models refuted by two types of preference reversals

Within the framework of prospect theory, various models for buying and selling prices of gambles can be constructed. However, two of these have already been rejected by considerable evidence, and it is worth noting why these theories do not work and have been rejected.

Let $b(g)$ and $s(g)$ represent the highest buying price and lowest selling price for gamble g, respectively (when gamble, g, is fixed, we use b and s, for simplicity).

First, one might theorize that people are willing to buy a gamble whenever $U(g) \geq U(b)$ and to sell whenever $U(s) \geq U(g)$. However, this model implies that if a person sets a higher price on gamble g than f then $U(g) > U(f)$, so the person should prefer g to f. However, there are gambles for which many individuals systematically set a higher selling or buying price on g than f and then prefer f over g in direct choice (Johnson & Busemeyer, 2005; Lichtenstein & Slovic, 1971; Mellers, Chang, Birnbaum, & Ordóñez, 1992; Mellers, Ordóñez, & Birnbaum, 1992; Tversky, Sattath, & Slovic, 1988).

Such results are called *preference reversals* because one way of comparing the utility of the gambles (direct choice) yields one preference relation, and another method (selling or buying price) yields an apparent contradiction. Furthermore, if buying and selling prices are each equal to the gamble's utility, then they should be equal to each other. Therefore, this first model can be rejected.

A second theory was proposed by Tversky and Kahneman (1991) for goods. Extended to gambles, the buying price is assumed to reflect an implicit comparison between the positive utility of receiving the gamble against the negative utility of (losing) the buying price, $U(-b)$. Similarly, selling price reflects an implicit comparison between the positive utility of receiving the sales price versus the loss of the gamble, $U(-g)$, where $U(-g)$ represents the utility of $-g = (-x, p; -y)$. In contrast to the first approach, this model involves comparisons between gains and losses, whereas the first approach involved only positive values. According to this theory,

$$U(g) + U(-b) = U(0); \quad (3)$$

$$U(s) + U(-g) = U(0). \quad (4)$$

Birnbaum and Zimmermann (1998, p. 176–178) proved that if we were to assume Eqs. (3) and (4), CPT (Eq. (1)) and loss aversion (Eq. (2)), and if $W^{+}(p) = W^{-}(p)$ (as assumed by Kahneman & Tversky, 1979 and reported as a good approximation by Tversky & Kahneman, 1992), it would follow that $U(s) = \lambda^2 U(b)$; therefore, $s = U^{-1}[U(b)\lambda^2]$; where U^{-1} is the inverse of $U(x)$. Thus, this model implies that selling and buying prices should be monotonically related to each other. Johnson and Busemeyer (2005) presented a similar proof, and concur that preference reversals between buying and selling prices in Birnbaum and Beeghley (1997) refute this (extended) model of Tversky and Kahneman (1991).

If in addition to Eqs. (1), (2), (3), and (4) we also assumed that $U(x) = x^{\beta}$, as in Tversky and Kahneman (1992); it would follow that $s = \lambda^{2/\beta} b$, so the ratio of selling price to buying price (WTA/WTP) should be a constant. Kahneman et al. (1990) listed empirical values of the ratio of WTA/WTP $= s/b$ from different studies, with a median slightly exceeding 4. If $\beta = 0.88$ (Tversky & Kahneman, 1992), then $s/b = 4$ implies $\lambda = 1.84$. Tversky and Kahneman (1991) realized that their theory implied that WTA/WTP of a \$5 bill must also be 4, so they postulated exceptions for cash or goods held for exchange.

More damaging than the need for such exceptions, data from several studies reported that selling prices and buying prices are not even monotonically related to each other, which contradicts this model (Birnbaum, 1982, p. 470–472; Birnbaum & Stegner, 1979). For example, Birnbaum and Sutton (1992) found that from the viewpoint of the seller, $g = (\$96, 0.5; \$0)$ is judged by the majority of individuals higher than $f = (\$48; 0.5; \$36)$; whereas, from the viewpoint of the buyer, g is judged lower than f by the majority of the same individuals. Such reversals between buying and selling prices represent a second type of preference reversal that has been found in several other studies (Birnbaum & Beeghley, 1997; Birnbaum et al., 1992; Birnbaum & Zimmermann, 1998). These preference reversals between WTA and WTP rule out this model of loss aversion and any other model in which the ratio, WTA/WTP, is constant.

Having ruled out these two models, we next take up another approach that is not rejected by these two types of preference reversals.

1.1.2. Third-generation prospect theory

This third approach, formulated by Birnbaum and Zimmermann (1998, p. 178–180) and independently by Schmidt et al. (2008), treats decisions to buy or sell as the result of an adjustment of the consequences of the gambles to reflect buying or selling prices. In the case of buying prices, the buyer is presumed to evaluate a new gamble, which we denote $g - b = (x - b, p; y - b)$, where $x - b$ is the profit if the gamble wins and $y - b$ is the loss if the gamble yields only y. Similarly, the seller considers it a "loss" of $x - s$ if the gamble might win x, since the seller gave up the opportunity to win; and if the gamble pays only y, the seller considers that a "win" of $s - y$. In decisions to sell, the integrated gamble is denoted as $s - g = (s - x, p; s - y)$; therefore,

$$U(g - b) = U(0); \quad (5)$$

$$U(s - g) = U(0). \quad (6)$$

Definition. Third-generation prospect theory (TGPT) consists of the following assumptions: Eqs. (5) and (6), CPT (Eq. (1)), Eq. (2) (loss aversion), and $U(x) = x^{\beta}$.

Theorem 1. *For gambles $g = (x, p;\ y)$ and $g' = (x, 1 - p; y)$, TGPT implies*

$$b(g) + s(g') = x + y, \quad (7)$$

where $b(g)$ and $s(g')$ represent the buying and selling prices for gamble g and g'.

Expression (7) is called *complementary symmetry*, even though the complementary gambles, g and $g\prime$ might be played independently on different trials. Intuitively, this property follows from the symmetry between buyers and sellers: a loss to one is a gain to the other.

Proof. From (5) and (2),and because $U(0) = 0$, highest buying prices of gamble, $g = (x, p; y)$, $x > y > 0$, are given by the following:

$$\begin{aligned} 0 &= W^{+}(p)U(x - b) + W^{-}(1 - p)U(y - b) \\ &= W^{+}(p)U(x - b) - W^{-}(1 - p)\lambda U(b - y). \end{aligned}$$

Define

$$T(p) = \left[\frac{W^{+}(p)}{W^{-}(1 - p)\lambda}\right]^{\frac{1}{\beta}},$$

then

$$b = \frac{T(p)x + y}{T(p) + 1}. \quad (8)$$

Similarly, for selling price of g,

$$s = \frac{x + T(1 - p)y}{T(1 - p) + 1}. \quad (9)$$

For $g = (x, p;\ y)$ and $g' = (x, 1 - p; y)$, it follows from Eqs. (8) and (9) that $b(g) + s(g') = x + y$, which proves Theorem 1. □

Note that the derivations of Eqs. (8), (9), and (7) did not assume anything about W^{+} and W^{-}, except that $T(p)$ is a number. Note that complementary symmetry (Eq. (7)) holds for all x, y, and p. The present study tests it by manipulating p, holding $x + y$ and $|x - y|$ fixed in each of several tests. The horizontal line in Fig. 1 depicts this prediction of TGPT. The other curves in Fig. 1 show that other theories (described in the next sections) violate this property, so we can compare theories by testing this property.[1]

TGPT does not imply that WTA/WTP is a constant. For example, suppose $W^{+} = W^{-}$, $U(x) = x$, and $\lambda = 2$. For $g = (\$100, 0.5;\ 0)$, it follows that $b = \$33$ and $s = \$67$, so $s/b = 2$. However, for $g = (x, p; y)$, when $x = y$ it follows that $s/b = 1$. Thus, this theory avoids the disproved implication of the theory of Tversky and Kahneman (1991), while retaining loss aversion.

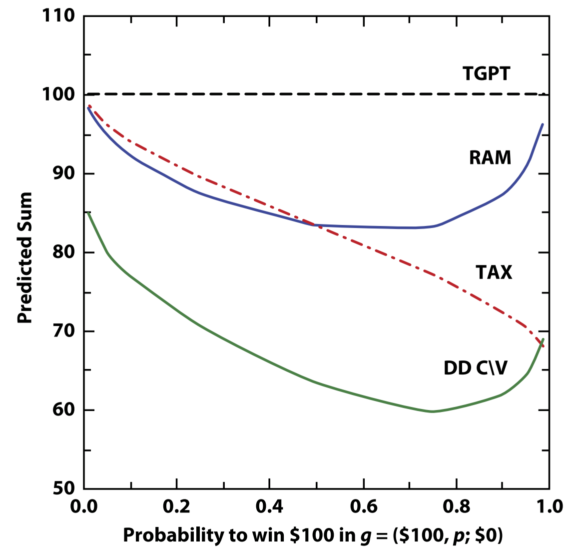

Fig. 1. Predictions of four models for test of complementary symmetry. Predicted sum of judgments of buying price of $g = (\$100, p; \$0)$ plus selling price of $g\prime = (\$100, 1 - p; \$0)$. Third-generation prospect theory (labeled TGPT) predicts a horizontal line equal to $x + y$, shown as the horizontal line at \$100. Duplex Decomposition C\V model (labeled DD C\V) predicts U-shape, as shown by the lowest curve. RAM model also predicts a U-shape with a minimum between 0.5 and 1. TAX model predicts a decrease to the right.

Definition (*Consequence Monotonicity (CM) for Binary Gambles*). $\forall\, g = (x, p;\ y)$, $\forall\, x^{+} > x$, and $\forall\, y^{+} > y$, $b(x^{+}, p;\ y) > b(g)$, $b(x, p;\ y^{+}) > b(g)$, $s(x^{+}, p;\ y) > s(g)$, and $s(x, p;\ y^{+}) > s(g)$.

Theorem 2. *TGPT implies CM in binary gambles.*

Proof. Eqs. (8) and (9) are both strictly increasing monotonic functions of x and of y. □

Definition. We say that g^{+} dominates g^{-} by *first order stochastic dominance* (FOSD) if the probability to win prize v or more in gamble g^{+} is always at least a high and sometimes higher than the probability to win v or more in gamble g^{-}, $\forall\, v$.

Definition (*Recipe of Birnbaum (1997, 2004)*). Let $g = (x, p;\ y)$, $g^{-} = (x, p - r; x^{-}, r; y, 1 - p)$, and $g^{+} = (x, p; y^{+}, t; y, 1 - p - t)$, where $x > x^{-} > y^{+} > y > 0$ and all of the probabilities are between 0 and 1.

For example, $g^{+} = (\$96, 0.9; \$14, 0.05; \$12, 0.05)$ dominates $g^{-} = (\$96, 0.85, \$90, 0.05; \$12, 0.1)$, because the probability to win \$96 or more is higher in g^{+} than in g^{-}, the probability to win \$90 or more is the same; the probability to win \$14 or more is higher in g^{+} than in g^{-} and the probability to win any other outcome or more is the same.[2]

[1] Michal Lewandowski (personal communication, April 23, 2016) reported that he has proved that complementary symmetry (7) holds for any monotonic utility function under (5) and (6); in which case tests of that property would not depend on the particular form of loss aversion assumed in (2).

[2] Birnbaum (1997) developed this recipe to illustrate that configural weight models (Birnbaum & Stegner, 1979) can predict violations of FOSD in such specially constructed choice problems. Birnbaum and Navarrete (1998) later reported that about 70% of undergraduates tested violated FOSD by choosing g^{-} over g^{+} in direct choices, and such violations have been replicated in a number of subsequent studies (e.g., Birnbaum, 2004). Because rank-dependent utility models (including CPT) must satisfy FOSD, such violations are strong evidence against the descriptive adequacy of those models. Birnbaum (2004) showed that the region of such predicted violations of FOSD by TAX is very small in the space of choices between three-branch gambles.

Definition. *Coalescing* is the assumption that if two branches of a gamble have identical consequences, they can be combined by adding their probabilities without affecting utility; e.g., $U(x, p; x, q; y, 1 - p - q) = U(x, p + q; z, 1 - p - q)$ and $U(x, p; y, q; y, 1 - p - q) = U(x, p; y, 1 - p)$.

Theorem 3. *TGPT implies that buying and selling prices satisfy FOSD in recipe of Birnbaum (1997, 2004).*

Proof. Birnbaum and Navarrete (1998, p. 57–58) proved that CPT implies coalescing for any decumulative probability weighting functions. They also proved (p. 53) that any theory satisfying transitivity, CM, and coalescing must satisfy FOSD in Birnbaum's recipe. Therefore, according to CPT, $U(g^+) > U(g) > U(g^-)$. By CM, $U(x-b, p-r; x^- -b, r; y-b) < U(x-b, p-r; x-b, r; y-b)$. From coalescing, $U(x-b, p-r; x^- -b, r; y-b) < U(x-b, p; y-b)$. Using coalescing (splitting) on the right side, we have $U(x-b, p-r; x^- - b, r; y-b) < U(x-b, p; y-b) = U(x-b; p; y-b, t; y-b)$; By CM, $U(x-b; p; y-b, t; y-b) < U(x-b; p; y^+ -b, t; y-b)$; by transitivity, $U(x-b, p-r; x^- -b, r; y-b) < U(x-b; p; y^+ -b, t; y-b)$; therefore (by Eq. (5)), $b(g^-) < b(g^+)$. The proof for selling prices works the same way, except $U(s - x, p - r; s - x^-, r; s - y) > U(s - x, p; s - y^+, t; s - y)$, so $s(g^-) < s(g^+)$ in Eq. (6). □

1.2. Joint receipts: rank- and sign-dependent utility

Luce (2000) deduced implications for buying and selling prices from his theory based on joint receipts of gains, losses, and gambles defined on those consequences. Joint receipt refers to psychological combination of two or more goods or gambles. One can view the act of buying as the joint receipt of a loss of money and of the gain of a gamble, and the act of selling as the joint receipt of cash and the loss of a gamble. Different combinations of assumptions in Luce's theoretical framework give rise to different models as special cases. Some of these models imply that buying and selling prices are identical, contrary to evidence, but others imply $s > b$. In contrast with loss aversion theories, however, the utility function for losses does not appear in the equations derived from this approach for gambles of positive consequences; instead, these models attribute the discrepancy between buying and selling prices to the pattern of weighting in mixed gambles.

This class of models, detailed in Luce (2000, Chapters 6–7), rests on two primitives in addition to a weakly-ordered (reflexive, transitive, and complete) preference relation, $\succsim$, over gambles. In the usual way, define indifference, $\sim$, strict preference, $\succ$, and the converse ordering, $\precsim$.

The first additional concept is a distinguished consequence, denoted e, that is identified as *no change from the status quo*. If f is a gamble (including, as a special case, pure consequences) and $f \succsim e$, then f is called a *gain* and if $f \precsim e$, then f is a *loss*.

The second additional concept is a binary operation, denoted $\oplus$ and called *joint receipt*. If f and g are two gambles (again, including pure consequences as special cases), then $f \oplus g$ means that the decision maker holds or receives them at the same time. We assume that $\oplus$ is commutative in the sense that $f \oplus g \sim g \oplus f$ and that e is its identity in the sense that $f \oplus e \sim f$. It is useful below to define a "subtraction" operation, $\ominus$, in terms of $\oplus$ by

$$f \ominus g \sim h \Longleftrightarrow f \sim g \oplus h. \tag{10}$$

We assume that such differences exist, but we do not assume that for money amounts x, y that $x \oplus y = x + y$. The reasons are discussed on pp. 144 and 154–155 of Luce (2000).

1.2.1. Buying and selling prices

The following models have some similarities to TGPT, in the sense that they replace actual subtraction with the generalized subtraction operator $\ominus$ of Expression (10), but the utility structure is quite different. Luce (1991; see also Hazen & Lee, 1991) postulated (maximum) buying and (minimum) selling prices, b and s, of a binary gamble $(x, p; y)$, with $x \succsim y$, as solutions to the following indifferences

$$(x \ominus b, p; y \ominus b) \sim e \quad \text{and} \quad (s \ominus x, p; s \ominus y) \sim e. \tag{11}$$

Note that if $\ominus \equiv -$, then this theory agrees with (5) and (6) of TGPT (Section 1.1.2).

The rationale is that if one buys the gamble for the amount b, then the net result is a gamble with the consequences reduced by the amount b. Similarly, if one sells a gamble for s, then one still "has" the gamble but with, in effect, the consequence of s less either x or y depending upon the outcome of the chance phenomena underlying the gamble. Note: from the assumptions that $\succsim$ is a weak order that satisfies consequence monotonicity over gambles, it follows that:

- $x \ominus b \succsim e \succsim y \ominus b$ and $s \ominus x \precsim e \precsim s \ominus y$.
- Thus, no matter whether the gamble is all gains, as it is in the present study, or all losses, or mixed gains and losses, the gambles defining the buying and selling prices necessarily have mixed consequences.
- For $b' \prec b$, $(x \ominus b', p; y \ominus b') \succ e$ and for $b' \succ b$, $(x \ominus b', p; y \ominus b') \prec e$.
- For $s' \succ s$, $(s' \ominus x, p; s' \ominus y) \succ e$ and for $s' \prec s$, $(s' \ominus x, p; s' \ominus y) \prec e$.

It is clear from the definitions we need to understand the form of the utility of gambles of a mixed gain and a loss. To do this entails two steps. The first is to decompose a mixed gamble into the joint receipt of two gambles that are not mixed ones. The second is to see the form of the utility of joint receipts. We take them up in that order.

1.2.2. Decomposing mixed gambles

For gambles of mixed gains and losses, two major hypotheses have been explored that are designed to reduce the calculation of $U(x, p; y)$, where $x \succsim e \succsim y$, to the joint receipt of gambles in which one of the consequences is e. Such simple gambles may be called *unitary*. The advantage of such decompositions is that if one can figure out how to calculate utility of a joint receipt, then the calculation is reduced to determining the form of U for unitary gambles. And for most models that have been proposed, they have the *separable* form

$$U(x, p; e) = U(x)W^{(i)}(p),$$
$$\text{where } i = + \text{ if } x \succsim e,\ i = - \text{ if } x \prec e.$$

Qualitative conditions for separability to hold are well understood (Luce, 2000, Section 3.5.2).

One decomposition assumption is *general segregation* (Luce, 1997), for $x \succsim y, x \succsim e$,

$$(x, p; y) \sim \begin{cases} (x \ominus y, p; e) \oplus y, & \text{if } (x, p; y) \succsim e \\ (e, p; y \ominus x) \oplus x, & \text{if } (x, p; y) \precsim e. \end{cases}$$

Note the normative quality of this condition in the sense that what one receives in the decomposed gamble is exactly the same as in the original gamble. This generalizes the property of segregation defined by Luce (1991) and Luce and Fishburn (1991) for gambles of all gains (or all losses).

The other proposed rule is *duplex decomposition* (Luce & Fishburn, 1991; Slovic & Lichtenstein, 1968)

$$(x, p; y) = (x, p; e) \oplus (e, p; y),$$

where the two unitary gambles on the right are run independently. Note that this is not normatively equivalent because the possible outcomes on the right are $x, y, x \oplus y$, and $e \oplus e = e$, whereas on the left they are just x or y. Nevertheless there are empirical data indicating that it may hold empirically (Luce, 2000, Section 6.2.3).[3]

1.2.3. Utility of joint receipt

So the task is reduced to determining the utility of joint receipt. Luce (2000, Section 4.4.6) developed the following formula on the assumption that $\oplus$ is a commutative operation and that segregation holds and f and g are gambles (or pure consequences) of both gains or both losses

$$U(f \oplus g) = U(f) + U(g) - \delta_i U(f)U(g),$$

where i distinguishes gains and losses. When $\delta_i = 0$, utility U, is additive. For $\delta_i \neq 0$, it is not difficult to show that, for gains and for losses separately, there exists an additive representation V, i.e.,

$$V(f \oplus g) = V(f) + V(g),$$

for which U is a negative exponential function of V (concave) when $\delta_i > 0$ and an exponential function (convex) when $\delta_i < 0$. This means that $\oplus$ is also associative over gains and losses separately.

Moreover, for money, invariance arguments imply that V for money is a power function (Luce, 2000, Section 4.5.3),

$$V(x) = \alpha x^{\beta}, \quad x > 0.$$

In the present context of buying and selling prices, negative consequences will play no direct role in the calculation, although they definitely affect the weights used and the forms of the equations derived. The cases where U itself is additive will be ignored since they result in equal buying and selling prices, which we know empirically is wrong. So there are four combinations with the gains first and losses second: concave\concave (denoted C\C), convex\convex (V\V), concave\convex (C\V), convex\concave (V\C).

1.2.4. Calculating the prices

Formulas for buying and selling prices were derived in Chapter 7 of Luce (2000, Section 7.4).[4] One arrives at equations for $U(b) = B$, $U(s) = S$ that are stated in terms of $U(f) = F$, $U(g) = G$, and various weighting functions over p or $1-p$ that differ depending on whether p applies to a gain or loss. The (normed) utility expressions for the concave cases are for the money amount $x > 0$

$$U(x) = 1 - \exp(-\kappa x^{\beta}), \tag{12}$$

and in the convex cases

$$U(x) = \exp(\kappa x^{\beta}) - 1. \tag{13}$$

For the weighting functions, we used the form derived by Prelec (1998) and deduced from a behavioral property by Luce (2001; see Luce, 2000, Section 3.4.3)

$$W^{(i)}(p) = \exp\left[-\gamma(i)\,(-\ln p)^{\eta(i)}\right], \tag{14}$$

where i is $+$ when the consequence associated with p is a gain and $-$ when it is a loss.

Thus, there are a total of 6 parameters for each of these models, two from the utility function of gains and 4 from the weights. And there are six models that do not lead to $b \sim s$, four based on duplex decomposition and two based on general segregation. Note that all of these require finding the "best" overall solution by varying the 6 parameters. We used minimization of sum of squares as the criterion of the fit.

In the duplex decomposition case the equations for B and S are quadratics, with exactly one real solution in the interval (G, F). For example, the equations for the duplex decomposition concave\convex (DD C\V) case are as follows:

$$\begin{aligned} 0 = {} & B^2 W^-(1-p) - B\left\{W^+(p) + W^-(1-p)\right. \\ & \left. - G[W^+(p) - W^-(1-p)]\right\} \\ & + FW^+(p) + GW^-(1-p) - FGW^+(p), \end{aligned} \tag{15}$$

$$\begin{aligned} 0 = {} & S^2 W^+(1-p) - S\left\{W^+(1-p) + W^-(p)\right. \\ & \left. + G[W^+(1-p) - W^-(p)]\right\} \\ & + FW^-(p) + GW^+(1-p) - FGW^-(p). \end{aligned} \tag{16}$$

1.3. Configural weighting and point of view

Birnbaum and Stegner (1979) theorized that buyers would place greater weight on the lower ranked estimates of value than do sellers. These rank-affected configural weight models correctly predicted that buyer's and seller's judgments of value are not monotonically related to each other (Birnbaum, 1982).

It has also been found that judgments are not always a monotonic function of the consequences (Birnbaum, 1997; Birnbaum et al., 1992; Birnbaum & Sutton, 1992). As the lowest consequence of a gamble is changed from zero to a small positive consequence, judgments can decrease. Such violations of CM can be described by the assumption that the zero consequence receives less weight than nonzero consequences.

For binary gambles of the form, $g = (x, p;\ y)$, where $x > y \geq 0$, RAM and TAX models are both relative weight averaging models of the form,

$$U[(x,p;y),j] = \begin{cases} U(x)W(p,j) + U(y)[1-W(p,j)] & y > 0 \\ U(x)W_0(p,j) + U(y)[1-W_0(p,j)] & y = 0 \end{cases} \quad (j = B, S).$$

Assuming $U(0) = 0$, the second equation can be simplified to $U[(x,p;0),j] = U(x)W_0(p,j)$ for $y = 0$. The relative weights of the more preferred consequence, $W(p,j)$ and $W_0(p,j)$, depend on two things: the probability, p, of that consequence x and the judge's viewpoint, $j = B$ or S (buyer or seller). In this approach, losses play no role in buying and selling prices of positive consequence gambles.

Utility is approximated as a power function of monetary prizes, $U(x) = x^{\beta}$.

In previous research, it has been found that the mapping from gambles' inverse utilities to overt responses can be approximated by a constant of proportionality, which is assumed to be the same for both viewpoints:

$$s = \nu U^{-1}[U(g,S)];$$
$$b = \nu U^{-1}[U(g,B)];$$

where $s(g)$ and $b(g)$ represent the selling and buying prices of gamble g; and ν is a constant (Birnbaum & Beeghley, 1997).

1.3.1. RAM model

In the Rank and Augmented sign-affected Multiplicative weights (RAM) model (Birnbaum, 1997), configural weights are represented as products of terms involving the effects of viewpoint, rank, augmented sign, and probability. Augmented sign has three levels, $+$, 0, and $-$.

3 However, Cho, Luce, and Truong (2002) empirically tested duplex decomposition and general segregation, and, depending upon the exact criterion used to say that a rule can be rejected, they found that from one third to one half the participants did not satisfy either.

4 Note that one of the relevant propositions in Luce (2000, 7.4.6(ii)), is in error. This was corrected in Birnbaum, Yeary, Luce, and Zhao (2002), and in an errata cited here for Luce (2000).

In the RAM model the weight of the higher valued, probability-consequence branch (to win x with probability p) is given by:

$$W(p,j) = \frac{\alpha_j p^\gamma}{\alpha_j p^\gamma + (1-p)^\gamma}, \quad j = B, S. \tag{17}$$

When $\gamma < 1$, this expression implies that the *relative* weight of the higher consequence will be an inverse-S function of p, even though the *absolute* weight of the higher consequence follows a power function of p. The inverse-S relationship between the estimated value of a binary gamble and probability has been reported by Gonzalez and Wu (1999), Tversky and Kahneman (1992), and others. The RAM model can be deduced from the assumption that the judge is estimating value so as to minimize an asymmetric loss function (Birnbaum et al., 1992; Weber, 1994).

The relationship between relative weights $W(p,j)$ and $W_0(p,j)$ can be fit with the following approximation:

$$W_0(p,j) = \alpha_0 W(p,j),$$

where α_0 is a constant reflecting greater relative weights for positive consequences when $y = 0$ relative to $y > 0$ ($\alpha_0 > 1$).

The three configural weight parameters are α_B, α_S, and α_0. The first two represent the relative weights of the branch with the higher valued consequence (relative to the branch with the lower-valued consequence) in buyer's and seller's viewpoints, respectively, and α_0, reflects the relative weight of the higher valued branch in the case of a lowest consequence of zero (relative to a positive lowest consequence in both viewpoints). The RAM model therefore has 6 parameters, ν, β, γ, α_B, α_S, and α_0.

1.3.2. TAX model

In the Transfer of Attention Exchange (TAX) model weights are transferred among probability-consequence branches and the sum of absolute weights is preserved. If there were no configural effects, the absolute weight of a branch would be a function of its probability, and its relative weight would be the ratio of this weight to total weight. However, in the buyer's viewpoint, weight is transferred from the branch with the higher-valued consequence and given to the lower consequence, whereas in the seller's viewpoint, weight is transferred in the opposite direction.

The relative weights of the branch to win x with probability p in the buyer's and seller's viewpoints are as follows:

$$W(p,B) = \frac{p^\gamma(1-\omega_B-\omega_0)}{p^\gamma + (1-p)^\gamma}, \tag{18}$$

$$W(p,S) = \frac{p^\gamma(1-\omega_0) + \omega_S(1-p)^\gamma}{p^\gamma + (1-p)^\gamma}. \tag{19}$$

The parameters, ω_B and ω_S, represent transfers of weight. In the buyer's viewpoint, ω_B, represents the proportion of weight of x (which would otherwise have been p^γ) that is taken from the higher ranked branch and given to the lower-ranked branch in the buyer's point of view. For sellers, ω_S, represents the proportion of the lower-ranked branch's weight [which would otherwise have been $(1-p)^\gamma$] transferred from the lower to the higher-valued branch. The parameter, ω_0 reflects weight transferred from the highest to lowest consequence in both viewpoints when $y > 0$; when $y = 0$, $\omega_0 = 0$. (In previous papers, TAX was written so that configural weights could be positive or negative, and the signs indicated whether weight was transferred from lower to higher consequences or vice versa. In this presentation, the directions of weight transfers are assumed to be known and the model is rewritten such that ω_B, and ω_S are both positive. The sign of ω_B must be reflected to compare it with values in previous publications.)

The TAX model thus has 6 free parameters, ν, β, γ, ω_B, ω_S, and ω_0. Previous research found that β in RAM and TAX can be set to 1 with little loss of predictive accuracy for $0 \le x < \$150$; RAM(5) and TAX(5) refer to RAM and TAX with β fixed to 1.

Table 1
Testable properties that distinguish the models. TGPT = Third generation prospect theory; DD C\V = Duplex Decomposition, Concave\Convex (joint receipt model); RAM = rank and augmented sign affected configural weight model with multiplicative weights; TAX = transfer of attention exchange model. The testable properties include CS = complementary symmetry; CM = consequence monotonicity; FOSD = first order stochastic dominance. TGPT implies satisfaction of CS, CM, FOSD; Viol = model does not satisfy the property for selected parameters. From parameters estimated in previous research, RAM and TAX predict violations for the stimuli in this study.

Model	Testable properties
TGPT	CS, CM, FOSD
DD C\V	Viol CS, Viol CM, Viol FOSD
RAM	Viol CS, Viol CM, Viol FOSD
TAX	Viol CS, Viol CM, Viol FOSD

1.4. Summary of testable properties

Besides comparing models based on an index of fit, models are evaluated by testing the properties of complementary symmetry, CM, and FOSD. Fig. 1 plots predictions for a test of complementary symmetry for four models: third-generation prospect theory (labeled TGPT), RAM, TAX, and the best-fitting of the joint receipt models (duplex decomposition with concave and convex utility for losses and gains, denoted DD C\V). Although TGPT implies complementary symmetry (horizontal line), the other models can violate it, as shown in Fig. 1.

TGPT implies that for any choice of probability weighting functions and any utility (value) functions, CM and FOSD must be satisfied. Luce (2000, p. 222–227) noted that violations of CM summarized by Birnbaum (1997) refute the models of buying and selling prices in Luce (2000, Chapter 6). However, Luce (2000, p. 259) stated that the DD C\V model might in principle violate CM (which would also violate first order stochastic dominance). Based on parameters estimated in previous studies, the configural weight models (RAM and TAX) predict violations of CM, violations of FOSD, and violations of complementary symmetry for the stimuli used in our design. These testable implications of the models are listed in Table 1.

2. Method

2.1. Instructions and stimuli

The judges were instructed to evaluate two- and three-outcome gambles from the viewpoints of buyer and seller. Gambles were displayed as in the following example

$$\frac{0.25\ 0.75}{\$40\ \$108}.$$

This display represents a binary gamble with a 0.25 probability to win \$40 and 0.75 probability to win \$108. Probability was described in terms of relative frequency by analogy to a can containing 100 otherwise identical slips that have the monetary consequences printed on them. Slips would be mixed, and one would be chosen blindly, at random, to determine the prize won. In this gamble, 25 slips say \$40 and 75 say \$108.

In the buyer's point of view, judges were asked to imagine that they were "deciding the most that a cautious buyer should pay to buy the chance to play the gamble". They were told that the buyer exchanges money for the chance to play the lottery. The sellers were asked to decide "the least that a seller should accept to sell the lottery". They were told the seller receives money and gives up the chance to play the lottery. Additional instructions were as in Birnbaum and Sutton (1992); these instructions produce smaller ratios of WTA/WTP, compared to studies reviewed in Horowitz and McConnell (2002) or Kahneman et al. (1990).

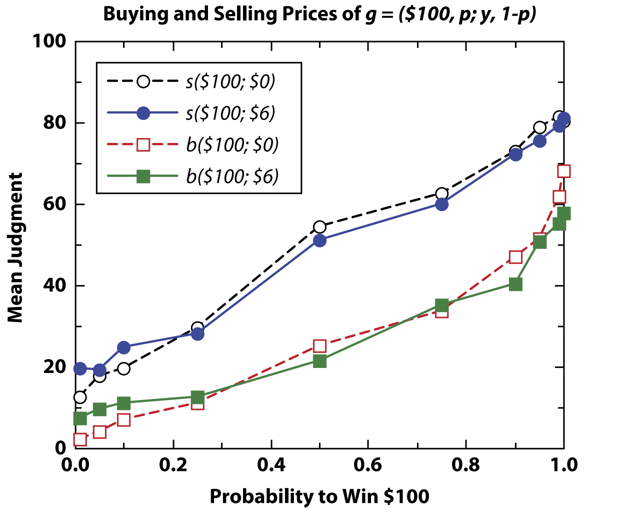

Fig. 2. Mean buying prices (WTP, squares) and selling prices (WTA, circles) of gambles of the form, ($100, p; y), are plotted as a function of p, with separate curves for $y = 0$ and $y = \$6$, averaged over participants. Filled symbols connected by solid curves ($y = \$6$) cross below the curves of open symbols connected by dashed curves ($y = \$0$) indicating violations of outcome monotonicity for large values of p.

2.2. Designs

There were 166 gambles, of which a subset of 63 binary gambles of the form $(y, p; x)$ are termed the *main* design of this article. These formed a 9 by 7 factorial design of p by (x, y) in which the 9 levels of p were 0.01, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, and 0.99, and the 7 levels of (x, y) were ($100, $0), ($72, $0), ($48, $0), ($24, $0), ($100, $6), ($100, $24) and ($100, $48). The *stochastic dominance* design consisted of 4 pairs of three-branch gambles. In Test 1, g^+ = ($96, 0.90; $14, 0.05; $12, 0.05) and g^- = ($96, 0.85; $90, 0.05; $12, 0.10); in Test 2, g^+ = ($97, 0.88; $5, 0.06; $3, 0.06), g^- = ($97, 0.84; $92, 0.04; $3, 0.12); in Test 3, g^+ = ($99, 0.95; $8, 0.03; $6, 0.02), g^- = ($99, 0.92; $91, 0.03; $6, 0.05); in Test 4, g^+ = ($97, 0.98; $7, 0.01; $4, 0.01), and g^- = ($97, 0.96; $89, 0.02; $4, 0.02); in each pair, g^+ dominates g^- by FOSD. These trials were interspersed randomly among the other trials.

2.3. Procedure

The 166 gambles were printed in restricted random orders in booklets, restricted so related gambles were not presented on consecutive trials. Each participant judged all 166 gambles from viewpoints of both buyer and seller. Buyers' and sellers' booklets were identical, with instructions for both points of view, except for the final paragraph in each booklet that reiterated the task at hand. Each booklet included 5 practice trials for the task and 6 additional unlabeled warm-up trials, preceding the 166 experimental trials. Half received the buyer's booklet first, followed by the seller's, and half received the other task order. Judges worked at their own paces and completed both tasks in two hours.

2.4. Judges

Judges (participants) were 66 undergraduates at California State University, Fullerton, who participated as one option toward an assignment in lower division psychology courses.

3. Results

3.1. Violations of Consequence Monotonicity (CM)

Violations of CM are shown in Fig. 2. The curves showing mean judgments of ($100, p; $0), averaged over participants, shown as open squares and circles connected by dashed lines cross the curves for ($100, p; $6), shown as filled circles and squares connected by solid lines. As reported previously (e.g., Birnbaum, 1992, 1997; Birnbaum et al., 1992 and Birnbaum & Sutton, 1992), increasing the lowest consequence from $y = 0$ to a small positive consequence increases judgments when the probability (p) of the highest consequence is small. However, when p is large, replacing $y = 0$ with a small positive value actually makes the judgment lower.

To test the significance of the crossover representing violations of CM, contrasts were constructed for each individual by computing the sum of six differences between judgments of ($100, p; $6) less those of ($100, p; $0). Summed over the three smallest values of p and over 2 viewpoints, 57 of 66 individuals had positive contrasts, indicating satisfaction of CM for small p. The mean contrast was $30.28, which is significantly greater than $0, $t(65) = 7.00$. However, at the highest three levels of p, the mean contrast was negative, −$19, which is significant, $t(65) = -3.33$ in the opposite direction. Of the 66 individuals, 40 showed negative contrasts compared to only 24 who had positive contrasts. Only 4 judges showed a crossover opposite to that shown in the means against 38 who showed both contrasts consistent with crossover in the means. Similarly, contrasts at $y = \$24$ (i.e., sum of differences in judgment between ($100, p; $24) and ($100, p; $0)) also showed the same crossover with a significant violation of consequence monotonicity for the three highest levels of p [$t(65) = -2.67$] and significant satisfaction for the three lowest levels of p, $t(65) = 16.11$. These violations of CM contradict the models of loss aversion and TGPT.

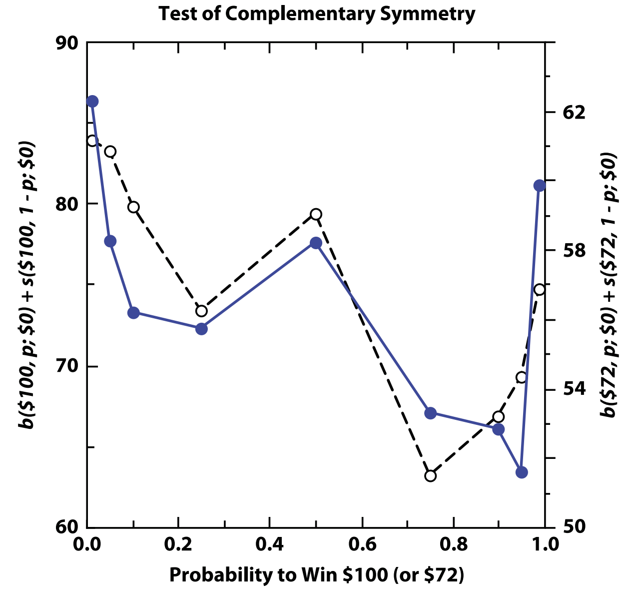

Fig. 3. Tests of complementary symmetry. According to third-generation prospect theory, sum of WTP + WTA of complementary gambles should be constant, independent of p. Filled circles show mean judgments of ($100, p; 0) plus its complement, averaged over participants; unfilled circles show judgments of ($72, p; 0) plus its complement (right ordinate). Compare to Fig. 1.

3.2. Violations of complementary symmetry

Eq. (7) (complementary symmetry) follows from TGPT. The sum of $b(x, p; y) + s(x, 1 - p; y)$ should be $x + y$, independent of p. In all 63 cells in the main design, the sum was less than $x + y$. In addition, the sum was not independent of p. Fig. 3 plots these sums for g = ($100, p; $0) and g = ($72, p; $0) as a function of p, averaged over participants, as in Fig. 1. Instead of being horizontal lines, the empirical curves have four features. First, the

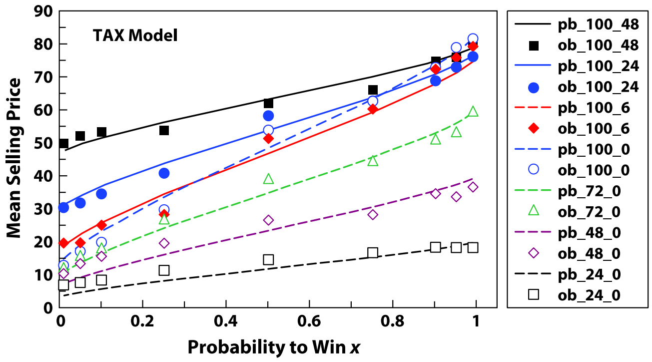

Fig. 4. Fit of TAX model with 5 free parameters to mean judgments of selling prices in the main design (WTA). Filled symbols show judgments of ($100, *p*; *y*), with squares, circles, and diamonds for $y = \$48$, $24, and $6, respectively. Solid curves show corresponding predictions (pb = predicted seller's price; ob = observed mean judgment, averaged over participants), as a function of probability to win $100. Unfilled symbols show judgments of gambles, (x, p; $0), with circles, triangles, diamonds, and squares for $x = \$100$, $72, $48, and $24, respectively, as a function of probability to win x. Dashed curves show corresponding predictions when $y = 0$. Cases where unfilled symbols exceed filled symbols prepresent empirical violations of consequence monotonicity. Cases where a dashed curve crosses solid curves show predicted violations of monotonicity.

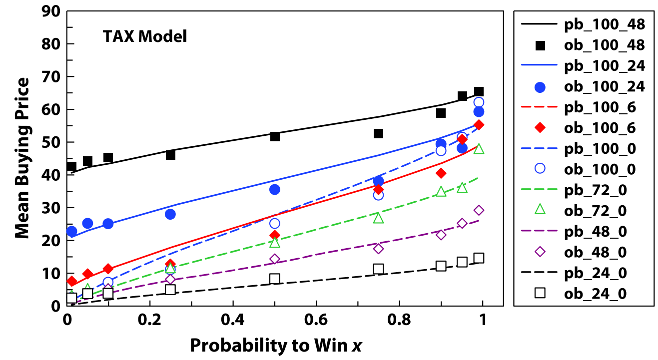

Fig. 5. Fit of the TAX model with 5 free parameters to mean judgments of buying prices (WTP), plotted as in Fig. 4 (pb = predicted buyer's price; ob = observed buyer's mean judgment).

Table 2
Comparison of fit of models to binary gambles.

Model (No. of parameters)	Sum of squared deviations
TGPT (9)	20,242
DD C\V(6)	8,822
DD C\V(12)	4,935
RAM (6)	1,129
RAM (5)	1,186
TAX(5)	1,051

sum of the judgments falls below $x + y$ ($100 or $72) for all p. Second, judgments decline as a function of p. Third, judgments show evidence of a U-shape, consistent with RAM and DD C\V models. Fourth, there is a "kink" at $p = 0.5$, not predicted by any of the models but similar to one reported by Birnbaum et al. (1992, p. 338–339).

A graph like Fig. 3 was drawn separately for each individual, and it was found that for each of the 66 participants, the sum $b(\$100, p; \$0) + s(\$100, 1 - p; \$0)$ was always less than $100; 47 individuals showed the decline, $b(\$100, 0.01; \$0) + s(\$100, 0.99; \$0) < b(\$100, 0.99; \$0) + s(\$100, 0.01; \$0)$; in 44 individuals the minimum of this sum fell in the interval, $0.5 < p < 0.99$. Thus, the main trends observed in the aggregate data of Fig. 3 are also characteristic of the majority of individual participants. In sum, these violations of complementary symmetry refute TGPT.

3.3. *Fit of models to binary gambles*

Four models were fit to the 126 mean judgments of the main design (63 binary gambles in 2 viewpoints) The sum of squared deviations for each model (and number of estimated parameters) is listed in Table 2.

Recall that the derivations of Eqs. (8) and (9) in TGPT did not assume anything about the forms of W^+ and W^- or their relationship. Because there are 9 levels of p in the data to be fit, there are 9 values of $T(p)$ to estimate. The least-squares estimates of $T(0.01)$, $T(0.05)$, $T(0.10)$, $T(0.25)$, $T(0.50)$, $T(0.75)$, $T(0.9)$, $T(0.95)$, $T(0.99)$ are 0.1373, 0.1865, 0.2277, 0.3523, 0.5556, 1.0422, 1.5099, 1.8304, and 2.4614. The sum of squared deviations for this model was 20,242.41, or a Root Mean Squared Deviation (RMS) = 12.67. TGPT fails badly to fit the data, despite using 9 free parameters. Forcing further restrictions on this model (e.g., adding assumptions concerning W^+ and W^-) would reduce the number of parameters, but make the fit even worse.

The best-fitting of the six joint-receipts models fit to the same aggregate data is the Duplex Decomposition model C\V (concave utility for gains and convex utility for losses). The least-squares parameter estimates are $\kappa = 0.2365$, $\beta = 0.5721$, $\gamma(+) = 0.6374$, $\eta(+) = 0.4498$, $\gamma(-) = 0.3161$, and $\eta(-) = 0.3665$; the sum of squared deviations is 8822.41, for a $RMS = 8.37$.

The parameter estimates for RAM(6) are $\nu = 0.844$, $\gamma = 0.505$, $\beta = 1.250$, $\alpha_B = 0.207$, and $\alpha_S = 0.890$, and $\alpha_0 = 1.100$. These six parameters yield a sum of squared deviations of 1128.65 or $RMS = 2.99$. When β is fixed to 1, RAM(5) model fits almost as well (1185.76), with best-fit parameter estimates of 0.823, 0.478, 1 (fixed), 0.306, 1.207, and 1.113, respectively.

TAX(5) fit the same data with smallest sum of squared errors: 1051, ($RMS = 2.89$). Parameter estimates, with $\beta = 1$ (fixed) are $\nu = 0.833$, $\gamma = 0.827$, $\omega_B = 0.336$, $\omega_S = 0.161$, and $\omega_0 = 0.092$. Allowing β to be free gave virtually the same parameter estimates and fit.

Clearly the two configural models fit best, the duplex decomposition one next, and TGPT fits least well.

The Duplex Decomposition models systematically failed to fit the relationship between judgments of gambles in which the lowest valued consequence is zero or positive. The best-fit solution did not reproduce the violations of CM (Fig. 2), as theorized it might do by Luce (2000, p. 259). To explore the source of deviations for the duplex decomposition models, the DD C\V model was fit to the mean judgments for these two subsets of the main design separately ($y = 0$ and $y > 0$), using 12 parameters. It was found that the sum of squared deviations for this version was 4934.48 ($RMS = 6.26$), which is still four times worse than the fit of either RAM or TAX with only 5 free parameters. The parameters for the utility function changed markedly between the two cases, which seems a complicated way to describe the data. Because models with 5 parameters fit better than models with 9 or 12 parameters, there is no need to do AIC analysis.

To check these conclusions for individuals, models were fit to every tenth participant, who were numbered from 1 to 66 in random order. Models were fit by function minimizing routines and the parameter space was also explored in a fine grid search to check for local minima. It was found that the order of the three categories of models was always the same for all seven individuals analyzed as it was for the averaged data. That is, configural weight models fit better than the best-fitting joint receipt model fit to individuals (which was for all individuals DD C\V), but there was no clear winner between the two configural models. And TGPT fits worst, despite having the largest number of free parameters.

If two models were equally accurate for fitting data of individual participants, the probability that one would beat the other for all seven individuals tested is $(\frac{1}{2})^7 = 0.008$.

Predictions for TAX(5) are shown as curves in Figs. 4 and 5, and mean judgments are shown as symbols. The TAX model

Table 3
Mean buying prices (WTP) in tests of stochastic dominance.

Test	g^+	g^-
1	$35	$54
2	$33	$47
3	$46	$53
4	$52	$58

Table 4
Mean selling prices (WTA) in tests of stochastic dominance.

Test	g^+	g^-
1	$65	$71
2	$64	$71
3	$74	$75
4	$73	$78

predicts violations of consequence monotonicity when lowest consequences change from $y = 0$ to $y > 0$ as p approaches 1. (Note the crossovers of the highest dashed curve and lowest solid curve and the corresponding crossover of open and filled symbols.) The model also does a fairly good job in describing the relationship between buying and selling prices.

Consistent with previous studies, the data in Figs. 4 and 5 show that the ratio of WTA/WTP is not a constant. The WTA/WTP ratio for binary gambles, $(x, p; y)$, varies inversely with p. The WTA/WTP ratios are 2.77, 2.07, 1.98, 1.89, 1.86, 1.59, 1.47, 1.44, 1.31, and 1.26, for the 9 levels of $p = .01$ to .99, respectively, averaged over $(x, y) =$ (\$100, \$0), (\$100, \$6), (\$100, \$24), and (\$100, \$48). Similarly, WTA/WTP varies inversely with $|x - y|$, averaged over the 9 levels of p; the mean ratios are 2.50, 1.93, 1.42, and 1.21 for $|x - y| =$ \$0, \$94, \$76, and \$52, respectively. The overall mean ratio was 1.91, smaller than the average values reported in reviews by Kahneman et al. (1990) and Horowitz and McConnell (2002). All four of the models compared here do not require that this ratio is constant; these variations in the ratio are best fit by the TAX and RAM models.

3.4. Violations of stochastic dominance (FOSD)

Tables 3 and 4 show that FOSD was violated in all 8 comparisons of mean judgments (4 tests and 2 viewpoints). The overall mean judgment was \$63.31 for dominated gambles, compared to \$55.11 for dominant gambles. This difference was significant, $F(1, 65) = 20.56$. The difference was greater for buyers (WTP), where the means were \$53.02 and \$41.18, than it was for sellers (WTA), where the means were \$73.61 and \$69.05, respectively. This interaction between viewpoint and dominance was significant, $F(1, 65) = 11.62$. Different gamble pairs (Tests, shown as Rows of Tables 3 and 4) showed significantly different magnitudes of violation, $F(3, 195) = 5.73$; but the three-way interaction of Tests by Dominance by Viewpoint was not significant, $F(3, 195) = 2.31$.

In separate analyses of each individual, it was found that 51 of 66 persons tested (77%) assigned higher mean judgments to dominated gambles than to corresponding dominant gambles, averaged over the four tests. Only 15 judges assigned higher mean judgments to the dominant gambles. Significantly more individuals violated than satisfied stochastic dominance in this recipe.

These results (Tables 3 and 4) agree with direct choices between the same pairs of gambles; Birnbaum and Navarrete (1998) found that 73, 61, 73, and 73 judges (out of 100) chose the dominated gamble g^- over the dominant, g^+, in direct choices in Tests 1 though 4, respectively. These violations are not consistent with TGPT or CPT, both of which imply satisfaction of FOSD in choice.

The data of Tables 3 and 4 were not used in the model-fitting, so they can be considered a type of hold-out sample to check the predictions of TAX(5). From the fit of two-branch gambles for buyers, $\omega_B + \omega_0 = 0.336 + 0.092 = 0.428$; in the TAX model of Birnbaum (2004), this means that $\delta/(n + 1) = -0.428$, where n is the number of branches (2 in this case), so $\delta_B = -1.28$. Similarly, for sellers, $\omega_S = 0.161$ and $\omega_0 = 0.092$ imply, $\delta_S = 0.207$. With parameters, $\nu = 0.833$, $\gamma = 0.827$, TAX predicts mean judgments of \$32 and \$50 for g^+ and g^- in WTP and \$71 and \$72 in WTA for the first test in Tables 3 and 4. Thus, TAX(5) with parameters estimated from the main design correctly predicts violations of FOSD. Parameters from previous studies also correctly predicted violations in this study (Birnbaum, 2004, p. 93; Birnbaum & Beeghley, 1997 and Birnbaum et al., 1992).

Violations of FOSD rule out TGPT and CPT, because there exist no parameters or monotonic value and weighting functions that allow those models to predict violations of FOSD.

4. Discussion

There are four major findings of the present experiment that are consistent with both of the configural weight models (RAM and TAX) but which are not consistent with at least one of the other models:

First, there are significant violations of complementary symmetry, which refute TGPT; these violations are apparent at both the aggregate level in the mean judgments of Fig. 3 and in analyses of individuals. RAM, TAX, and DD C\V can predict violations of complementary symmetry (Fig. 1).

Second, there are significant violations of consequence monotonicity at both aggregate and individual level, which refute TGPT. Both configural weight models can describe violations of consequence monotonicity where increasing a consequence from 0 to a small positive value can decrease the lottery's judged value (Fig. 2). The best-fitting of the joint receipt models failed to reproduce this aspect of the data.

Third, violations of first order stochastic dominance in both buying and selling prices refute TGPT, but were correctly predicted by TAX model using parameters estimated either from previous studies or from other data in this study. The fact that violations of first order stochastic dominance are found at both aggregate and individual level, for buying prices (Table 3), selling prices (Table 4), and choices (Birnbaum & Navarrete, 1998) indicates that these violations are not due to something unique to either the choice task or to judgment, but instead likely due to a common evaluation mechanism.

Fourth, both configural weight models fit the data for binary gambles better than the other models, despite using fewer parameters. No clear winner between TAX and RAM was determined; RAM appears to better approximate the violations of complementary symmetry, but TAX fit slightly better overall. The sum of squared deviations for TGPT with 9 parameters was more than 19 times larger than that for TAX with 5 parameters. The best-fitting joint receipt model (DD C\V) even when allowed 12 parameters had 4 times the sum of squares of TAX with 5 parameters.

From the fit of the models derived from joint receipts, it appears that violations of consequence monotonicity reported in previous judgment studies as well as in this study are not well-described by the properties of general segregation or duplex decomposition in additive joint receipt models. Recall that Cho et al. (2002) found that from one third to one half the participants did not satisfy either general segregation or duplex decomposition. This means that the joint receipt models discussed in this paper may very well be founded on wrong premises. For the present data, the model failed to adequately describe the fact that people treat the case with \$0 as a consequence differently from cases with strictly positive consequences, even when different sets of parameters were allowed in these two cases. This finding causes trouble for the previously published models derived in Luce's (2000, Chapter 7) approach because they are based on decomposing a general gamble to unitary ones (i.e., with one consequence $0), as in general segregation and duplex decomposition. Some progress has been made on this issue in Luce (2003, 2010).

These results are compatible with other research testing implications of configural weight models for buying and selling prices (Birnbaum, 1982; Birnbaum & Beeghley, 1997; Birnbaum et al., 1992; Birnbaum & Stegner, 1979; Birnbaum & Sutton, 1992; Birnbaum & Veira, 1998; Birnbaum & Zimmermann, 1998; Johnson & Busemeyer, 2005). Ashby, Dickert, and Glöckner (2012) found that buyers spend more time looking at the lower valued consequences than do sellers at all levels of probability, as if time spent looking at a component is an indicator of its configural weight.

Is there some other way to salvage CPT and "loss aversion"? It might seem that if we were to allow the parameters of CPT to be contingent on the judge's point of view, then we might be able to salvage that model, as follows:

$$b = J_b[U_b(x, p; y)],$$
$$s = J_s[U_s(-x, p; -y)],$$

where J_b and J_s represent monotonic functions that might reflect different "biases" produced by the request to produce numerical judgments in the buying and selling tasks, respectively; U_b and U_s represent utilities (values) for the gambles, allowing functions, W_b^+, W_s^+, W_b^-, W_s^-, U_b, and U_s that differ for buyer or seller. This model would allow any relationship between $U(x)$ and $U(-x)$, generalizing loss aversion. Although this model uses many parameters and can be made equivalent to RAM or TAX for certain binary gambles (with suitable choices of functions and parameters), this model still must satisfy consequence monotonicity and FOSD, so it is refuted by violations of those critical properties, despite the large number of free parameters and functions.

At one time it was thought that CPT might be an accurate description of choices among risky prospects. However, CPT implies properties such as first order stochastic dominance that are systematically violated by empirical data, whether analyzed at the aggregate or individual participant level (Birnbaum & Bahra, 2012; Birnbaum & Navarrete, 1998).

Marley and Luce (2005) re-summarized data that had been summarized in previous publications by Birnbaum and his colleagues, who had devised a number of critical tests of CPT such as upper and lower cumulative independence and had shown that these critical properties were systematically violated by aggregate data. Luce and Marley (2005) axiomatized a rank weighted utility representation that has some of the features of the configural weight models and generalizes CPT; see correction in Marley and Regenwetter (in press). The strength of the case against CPT continues to grow, as shown in reviews by Birnbaum (2008) and Birnbaum and Bahra (2012), who analyzed each individual's data using a model in which each choice problem can have a different error rate.

As found in previous studies, the "gap" between WTA and WTP is not a constant. A good deal of research has been based on the theory that the ratio, WTA/WTP, is a constant, and many studies estimated this ratio with a single item, such as a mug. With a single item, it is not possible to test the theory that a ratio is constant. Within this branch of the literature, rival theories arose:

According to the loss aversion theory of Tversky and Kahneman (1991, 1992), the "correct" ratio of WTA/WTP is about 4, and it can be calculated from the loss aversion parameter (λ). Variations of this theory were proposed, for example, in which buying prices were free of loss aversion, in which case the predicted ratio would be about 2. According to classical EU theory, the gap should be quite small for persons who are not poor, so the "correct" ratio of WTA/WTP should be close to 1. According to these theories the ratio is a constant, and many studies were devoted to estimating the ratio and comparing ratios between articles (Erickson et al., 2014; Horowitz & McConnell, 2002; Kahneman et al., 1990; Plott & Zeiler, 2005).

Some who argued for one value or the other also argued that the "right" way to do the study is the way that leads to the value consistent with one theory or the other. However, when theory is in doubt, it becomes circular to choose procedures based on how the results conform with one theory or the other. Instead, we think that proper theory of WTA and WTP will recognize that the gap is not a constant. In addition, we think the best theory should be able to predict the size of the gap based on details of the stimuli, experimental instructions, incentives, and procedures.

According to Birnbaum and Stegner's (1979) theory, configural weights can be manipulated by factors that affect a person's costs of over- as opposed to under-estimation (see also Birnbaum et al., 1992). In these models, it is the configural weights that produce risk aversion or risk seeking. The risk averse person transfers weight to the worst outcomes, features, or estimates of value of a gamble or object, whereas the risk-seeking person puts greater configural weight on higher outcomes, aspects or estimates.

According to the concept of endowment, there are just two levels of endowment for a given item: one either owns something or not. Thus, there are only two possible viewpoints one can use in a study: buyer and seller. However, as noted in Birnbaum and Stegner (1979), there are many possible viewpoints; for example, participants can be asked to judge the "true" or "fair" price of an object. For example, person A destroys person B's property, and a neutral judge is asked to judge the "fair" value of the object such that it would be "just and fair" for person A to pay B in compensation. Participants have no difficulty understanding this task, which is performed every day by judges and juries.

To observe the so-called "endowment" effect, one need not endow anyone with anything. Simply ask participants to identify with buyer, seller, or neutral. From the perspective of viewpoint theory, it should be possible to study further differentiation by asking people to advise a "conservative" or "cautious" buyer, a "moderate" buyer, or a "venturesome" buyer. Indeed, financial analysts devise portfolios for such buyers of different viewpoints with different proportions of stocks, bonds, and other investments.

Some economists might say that transfers of money in the courts and advice given to investors are outside the realm of economic theory; but these topics are not outside the realm of the psychological theories compared here. In other judgment domains, viewpoints also differ: in a court trial, there are defendant, prosecution, judge, and jury; in politics, there are candidates who are republicans or democrats, and voters who might also be republican or democrat. It seems reasonable that a common cognitive mechanism has evolved in humans to handle these analogous tasks.

With respect to the present experiment, one can reasonably ask if results might change with different procedures, incentives, or instructions. For example, these data represent judgments of buying and selling prices. Would the same conclusions be reached with choice-based prices? Similarly, might these results be altered if differently worded instructions or different incentives had been used? If so, then we need to develop theories of how results depend on these factors. According the configural weight models, instructions or incentives that affect the costs of over- or under-estimation should affect configural weights.

The estimated parameters in the present study for undergraduates show strong risk aversion in the buyer's viewpoint and are closer to risk neutral in the seller's. Perhaps these configural weights result from greater experience by undergraduates in buying rather than selling.

5. Conclusions

Of the four models of buying and selling prices evaluated in this study, the two oldest (RAM and TAX) were more accurate in predicting results than models by Luce (2000) based on his theory of joint receipts. Systematic violations of complementary symmetry, consequence monotonicity, and first order stochastic dominance rule out third generation prospect theory, which fit least well at both aggregate and individual level. In RAM, TAX, and Luce's models, the utility function for negative consequences plays no role in buying and selling prices of gambles with non-negative consequences; instead, the gap between WTP and WTA ("endowment" effects) are explained by buyers and sellers assigning different configural weights to lower and higher consequences of a gamble or aspects of a good. Our results for WTP and WTA can be explained without postulating loss aversion.

Acknowledgments

R. Duncan Luce passed away in 2012, and could therefore not contribute to the most recent versions of this paper. He developed the theory and models in Section 1.2 specifically to describe these data and he wrote that section as well as the results and discussion associated with it; he supervised the work of Li Zhao, who fit the models to individual subjects, and he made numerous other contributions to earlier versions of this paper.

This research was supported by National Science Foundation Grants, SBR-9410572 and SES 9986436 to C.S.U.F. and SBR-9808057 to UCI.

References

Ashby, N. J. S., Dickert, S., & Glöckner, A. (2012). Focusing on what you own: Biased information uptake due to ownership. *Judgment and Decision Making*, *7*, 254–267.

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

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science*, *3*, 310–314.

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

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

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

Birnbaum, M. H., & Bahra, J. P. (2007). Gain-loss separability and coalescing in risky decision making. *Management Science*, *53*, 1016–1028.

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

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

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

Birnbaum, M. H., & Navarrete, J. (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., & Veira, R. (1998). Configural weighting in judgments of two- and four-outcome gambles. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 216–226.

Birnbaum, M. H., Yeary, S., Luce, R. D., & Zhao, L. (2002). *Empirical evaluation of theories for buying and selling prices of binary gambles. Institute for Mathematical Behavioral Sciences Technical Report Series, MBS 02-12*.

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*(2), 145–187.

Cho, Y.-H., Luce, R. D., & Truong, L. (2002). Duplex decomposition and general segregation of lotteries of a gain and a loss: An empirical evaluation. *Organizational Behavior and Human Decision Processes*, *89*, 1176–1193.

Coombs, C. H., Bezembinder, T. G., & Goode, F. M. (1967). Testing expectation theories of decision making without measuring utility or subjective probability. *Journal of Mathematical Psychology*, *4*, 72–103.

Erickson, K. M., Marzilli, , & Fuster, A. (2014). The endowment effect. *Annual Review of Economics*, *6*, 555–579.

Gonzalez, R., & Wu, G. (1999). On the shape of probability weighting function. *Cognitive Psychology*, *38*, 129–166.

Hazen, G. B., & Lee, J.-S. (1991). Ambiguity aversion in the small and in the large for weighted linear utility. *Journal of Risk and Uncertainty*, *4*, 177–212.

Horowitz, J. K., & McConnell, K. E. (2002). A review of WTA/WTP Studies. *Journal of Environmental Economics and Management*, *44*, 426–447.

Irwin, J., Slovic, P., Lichtenstein, S., & McClelland, G. (1993). Preference reversals and the measurement of environmental values. *Journal of Risk and Uncertainty*, *6*, 5–18.

Johnson, J. G., & Busemeyer, J. R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychological Review*, *112*, 841–861.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the coarse theorem. *Journal of Political Economy*, *98*, 1325–1348.

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

Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics*, *99*, 507–521.

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

Luce, R. D. (1991). Rank- and sign-dependent linear utility models for binary gambles. *Journal of Economic Theory*, *53*, 75–100.

Luce, R. D. (1997). Associative joint receipts. *Mathematical Social Sciences*, *34*, 51–74.

Luce, R. D. (2000). *Utility of gains and losses: measurement-theoretical and experimental approaches*. Mahwah, NJ: Lawrence Erlbaum Associates, Errata: see Luce Web page at http://www.imbs.uci.edu/files/personnel/luce/luce.html.

Luce, R. D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology*, *45*, 167–179.

Luce, R. D. (2003). Increasing increment generalizations of rank-dependent theories. *Theory and Decision*, *55*, 87–146.

Luce, R. D. (2010). Behavioral assumptions for a class of utility theories: A program of experiments. *Journal of Risk and Uncertainty*, *41*, 19–27.

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., & Marley, A. A. J. (2005). Ranked additive utility representations of gambles: Old and new axiomatizations. *Journal of Risk and Uncertainty*, *30*, 21–62.

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

Marley, A.A.J., & Regenwetter, M. (in press). Choice, Preference, and Utility: Probabilistic and Deterministic Representations. In W. Batchelder, H. Colonius, E. Dzhafarov, J. Myung (Eds.), *New handbook of mathematical psychology. Volume 1: measurement and methodology.* Cambridge: Cambridge University Press.

Mellers, B. A., Chang, S., Birnbaum, M. H., & Ordóñez, L. (1992). Preferences, prices, and ratings in risky decision making. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 347–361.

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.

Morewedge, C. K., & Giblin, C. E. (2015). Explanations of the endowment effect: An integrative review. *Trends in Cognitive Sciences*, *19*, 339–348.

Plott, C. R., & Zeiler, K. (2005). The willingness to pay–willingness to accept gap, the "endowment effect", subject misconceptions, and experimental procedures for eliciting valuations. *American Economic Review*, *95*, 530–545.

Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.

Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty*, *36*, 203–223.

Slovic, P., & Lichtenstein, S. (1968). Importance of variance preferences in gambling decisions. *Journal of Experimental Psychology*, *78*, 646–654.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, *1*, 39–60.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, *106*(4), 1039–1061.

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.

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