CHAPTER 14

Belief and Preference in Decision Under Uncertainty*

Craig R. Fox
Anderson School of Management and Department of Psychology, UCLA, USA
and
Kelly E. See
Duke University, Fuqua School of Business, USA

14.1 INTRODUCTION

Most decisions in life are gambles. Should I speed up or slow down as I approach the yellow traffic light ahead? Should I invest in the stock market or in treasury bills? Should I undergo surgery or radiation therapy to treat my tumor? From mundane choices rendered with scarcely a moment's reflection to urgent decisions founded on careful deliberation, we seldom know in advance and with certainty what the consequences of our choices will be. Thus, most decisions require not only an assessment of the attractiveness of potential consequences, but also some appraisal of their likelihood of occurrence.

Virtually all decision theorists agree that values and beliefs jointly influence willingness to act under uncertainty. However, there is considerable disagreement about how to measure values and beliefs, and how to model their influence on decisions. Our purpose in this chapter is to bring into sharper focus the role of values and beliefs in decision under uncertainty and contrast some recent developments in the descriptive modeling of choice under uncertainty with the classical normative model.

14.1.1 The Classical Theory and the Sure-Thing Principle

The primitives of most decision theories are acts, states, and consequences (Savage, 1954; for an alternative approach, see Luce, 2000). An *act* is an action or option that yields one

* The view of decision making under uncertainty outlined in this chapter was heavily influenced by the late Amos Tversky. Of course, any errors or deficiencies of the present work are entirely the responsibility of the authors. We thank Jim Bettman, Rick Larrick, Bob Nau, John Payne, Shlomi Sher, Peter Wakker and George Wu for helpful comments on earlier drafts of this chapter. Most of the work on this chapter was completed while Craig Fox was at the Fuqua School of Business, Duke University, whose support is gratefully acknowledged.

Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making. Edited by David Hardman and Laura Macchi.

Table 14.1 A decision matrix. Columns are interpreted as states of the world, and rows are interpreted as acts; each cell entry x_{ij} is the consequence of act i if state j obtains

		States				
		s_1	...	s_j	...	s_n
A	a_1	x_{11}	...	x_{1j}	...	x_{1n}
C	...	...	...	...	...	...
T	a_i	x_{i1}	...	x_{ij}	...	x_{in}
S	...	...	...	...	...	...
	a_m	x_{m1}	...	x_{mj}	...	x_{mn}

of a set of possible *consequences* depending on which future *state* of the world obtains. For instance, suppose I am considering whether or not to carry an umbrella. Two possible acts are available to me: carry an umbrella, or do not carry an umbrella. Two relevant states of the world are possible: rain or no rain. The consequence of the act that I choose is a function of both the act chosen (which governs whether or not I am burdened by carrying an umbrella) and the state that obtains (which influences whether or not I will get wet).

More formally, let S be the set of possible states of the world, subsets of which are called events. It is assumed that exactly one state obtains, which is unknown to the decision maker. Let X be a set of possible consequences (also called "outcomes"), such as dollars gained or lost relative to the status quo. Let A be the set of possible acts, which are interpreted as functions, mapping states to consequences. Thus, for act $a_i \in A$, state $s_j \in S$, and consequence $x_{ij} \in X$, we have $a_i(s_j) = x_{ij}$. This scheme can be neatly captured by a decision matrix, as depicted in Table 14.1.

In the classical normative model of decision under uncertainty, decision makers weight the perceived attractiveness (utility) of each potential consequence by its perceived likelihood (subjective probability). Formally, if $u(x_{ij})$ is the utility of outcome x_{ij} and $p(s_j)$ is the subjective probability that state s_j will obtain, then the decision maker chooses the act that maximizes subjective expected utility (SEU):

$$SEU(a_i) = \sum_{j=1}^{n} u(x_{ij})p(s_j). \tag{1}$$

Hence, the classical model segregates belief (probability) from value (utility). Subjective expected utility theory (Savage, 1954) articulates a set of axioms that are necessary and sufficient for the representation above, allowing subjective probability and utility to be measured simultaneously from observed preferences.[1] For instance, if Alan is indifferent between receiving \$100 if it rains tomorrow (and nothing otherwise) or \$100 if a fair coin lands heads (and nothing otherwise), then we infer that he considers these target events to be equally likely (that is, $p(\text{rain}) = p(\text{heads}) = 1/2$). If he is indifferent between receiving one of these prospects or \$35 for sure, we infer that $u(35) = 1/2\, u(100)$. It is important to emphasize that Savage, following the tradition of previous theorists (e.g., Borel, 1924; Ramsey, 1931;

[1] For alternative axiomatic approaches that more explicitly distinguish the role of objective versus subjective probabilities, see Anscombe and Aumann (1963) and Pratt, Raiffa and Schlaifer (1964).

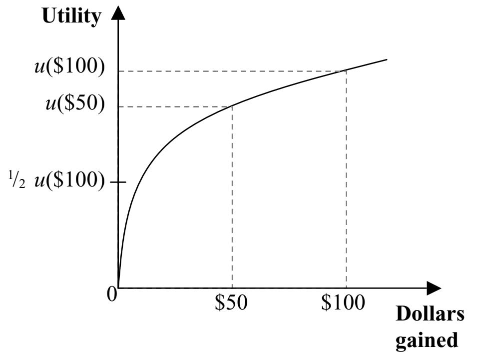

Figure 14.1 A concave utility function for dollars gained set to $u(0) = 0$

de Finetti, 1937), rejected direct judgments of likelihood in favor of a measure derived from observed preferences. In contrast, psychologists (e.g., Kahneman, Slovic & Tversky, 1982) give credence to direct expressions of belief and assume that they can be used to predict willingness to act under uncertainty.[2]

Expected utility theory was originally developed to explain attitudes toward risk. The lay concept of "risk" entails the threat of harm or loss. For instance, managers see risk as increasing with the likelihood and magnitude of potential losses (e.g., March & Shapira, 1987). Decision theorists, in contrast, see risk as increasing with variance in the probability distribution of possible outcomes, regardless of whether or not a potential loss is involved. For instance, a prospect that offers a .5 chance of receiving $100 and a .5 chance of receiving nothing is more risky than a prospect that offers $50 for sure—even though the "risky" prospect entails no possibility of losing money. *Risk aversion* is defined by decision theorists as a preference for a sure outcome over a chance prospect with equal or greater expected value.[3] Thus, the preference of $50 for sure over a 50–50 chance of receiving $100 or nothing is an expression of risk aversion. *Risk seeking*, in contrast, is defined as a preference for a chance prospect over a sure outcome of equal or greater expected value. It is commonly assumed that people are risk averse, and this is explained in expected utility theory by a concave utility function (see Figure 14.1). Such a shape implies, for example, that the utility gained from receiving $50 is more than half the utility gained from receiving $100; hence, receiving $50 for sure is more attractive than a 50–50 chance of receiving $100 or nothing.[4]

As stated earlier, Savage (1954) identified a set of preference conditions that are both necessary and sufficient to represent a decision maker's choices by the maximization of subjective expected utility. Central to the SEU representation (Equation (1)) is an axiom known as the "sure-thing principle" (also sometimes referred to as "weak independence"): if two acts yield the same consequence when a particular state obtains, then preference between acts should not depend on the particular nature of that common consequence

[2] Some statisticians, philosophers, and economists have been sympathetic to the use of direct probability judgment as a primitive for decision theories (e.g., DeGroot, 1970; Shafer, 1986; Karni & Mongin, 2000).

[3] The expected value of a gamble that pays $$x$ with probability p is given by xp. This is the mean payoff that would be realized if the gamble were played an infinite number of times.

[4] In expected utility theory, utility is a function of the decision maker's aggregate wealth and is unique up to a positive affine transformation (that is, utility is measured on an interval scale). To simplify (and without loss of generality), we have set the utility of the present state of wealth, $u(W_0) = 0$.

(see Savage, 1954). To illustrate, consider a game in which a coin is flipped to determine which fruit Alan will receive with his lunch. Suppose that Alan would rather receive an apple if a fair coin lands heads and a *cantaloupe* if it lands tails (a, H; c, T) than receive a banana if the coin lands heads and a *cantaloupe* if it lands tails (b, H; c, T). If this is the case, Alan should also prefer to receive an apple if the coin lands heads and *dates* if the coin lands tails (a, H; d, T) to a banana if it lands heads and *dates* if it lands tails (b, H; d, T). In fact, the preference ordering over these prospects should not be affected at all by the nature of the common consequence—be it a cantaloupe, dates or a subscription to *Sports Illustrated.* The sure-thing principle is necessary to establish a subjective probability measure that is additive (that is, $p(s_1) + p(s_2) = p(s_1 \cup s_2)$).[5]

14.1.2 Violations of the Sure-Thing Principle

The sure-thing principle seems on the surface to be quite reasonable, if not unassailable. In fact, Savage (1954, p. 21) wrote, "I know of no other extralogical principle governing decisions that finds such ready acceptance." Nevertheless, it was not long before the descriptive validity of this axiom was called into question. Notably, two "paradoxes" emerged, due to Allais (1953) and Ellsberg (1961). These paradoxes pointed to deficiencies of the classical theory that have since given rise to a more descriptively valid account of decision under uncertainty.

Problem 1: The Allais Paradox

Choose between: (A) \$1 million for sure; (B) a 10 percent chance of winning \$5 million, an 89 percent chance of winning \$1 million, and a 1 percent chance of winning nothing.

Choose between: (C) an 11 percent chance of winning \$1 million; (D) a 10 percent chance of winning \$5 million.

Problem 2: The Ellsberg Paradox

An urn contains 30 red balls, as well as 60 balls that are each either white or blue (but you do not know how many of these balls are white and how many are blue). You are asked to draw a ball from the urn without looking.

Choose between: (E) win \$100 if the ball drawn is red (and nothing if it is white or blue); (F) win \$100 if the ball drawn is white (and nothing if it is red or blue).

Choose between: (G) win \$100 if the ball drawn is either red or blue (and nothing if it is white); (H) win \$100 if the ball drawn is either white or blue (and nothing if it is red).

Maurice Allais (1953) presented a version of Problem 1 at an international colloquium on risk that was attended by several of the most eminent economists of the day. The majority of

[5] Savage's postulates P3 and P4 establish that events can be ordered by their impact on preferences and that values of consequences are independent of the particular events under which they obtain. In the context of P3 and P4, the sure-thing principle (Savage's P2) establishes that the impact of the events is additive (that is, that it can be represented by a probability measure). For an alternative approach to the derivation of subjective probabilities that neither implies nor assumes expected utility theory, see Machina and Schmeidler (1992).

Table 14.2 Visual representation of the Allais paradox. People typically prefer *A* over *B* but also prefer *D* over *C*, in violation of the sure-thing principle

	Ticket numbers		
	1	2–11	12–100
A	\$1M	\$1M	\$1M
B	0	\$5M	\$1M
C	\$1M	\$1M	0
D	0	\$5M	0

Table 14.3 Visual representation of the Ellsberg paradox. People typically prefer *E* over *F* but also prefer *H* over *G*, in violation of the sure-thing principle

	30 balls	60 balls	
	Red	White	Blue
E	\$100	0	0
F	0	\$100	0
G	\$100	0	\$100
H	0	\$100	\$100

these participants favored (A) over (B) and (D) over (C). Daniel Ellsberg (1961) presented a version of Problem 2 as a thought experiment to many of his colleagues, including some prominent decision theorists. Most favored (E) over (F) and (H) over (G). Both of these patterns violate the sure-thing principle, as can be seen in Tables 14.2 and 14.3.

Table 14.2 depicts the Allais problem as a lottery with 100 consecutively numbered tickets, where columns denote ticket numbers (states) and rows denote acts. Table entries indicate consequences for each act if the relevant state obtains. It is easy to see from this table that for tickets 1–11, acts *C* and *D* yield consequences that are identical to acts *A* and *B*, respectively. It is also easy to see that for tickets 12–100, acts *A* and *B* yield a common consequence (receive \$1 million) and acts *C* and *D* yield a different common consequence (receive nothing). Hence, the sure-thing principle requires a person to choose *C* over *D* if and only if she chooses *A* over *B*. The modal preferences of *A* over *B* and *D* over *C*, therefore, violate this axiom.

Table 14.3 depicts the Ellsberg problem in a similar manner. Again, it is easy to see that the sure-thing principle requires a person to choose *E* over *F* if and only if he chooses *G* over *H*. Hence the dominant responses of *E* over *F* and *H* over *G* violate this axiom.

It should also be apparent that the sure-thing principle is implied by expected utility theory (see Equation (1)). To illustrate, consider the Allais problem above (see Table 14.2), and set $u(0) = 0$. We get:

$$SEU(A) = .11u(\$1\text{M}) + .89u(\$1\text{M})$$
$$SEU(B) = .10u(\$5\text{M}) + .89u(\$1\text{M}).$$

Similarly,

$SEU(C) = .11u(\$1\text{M})$
$SEU(D) = .10u(\$5\text{M}).$

Hence, the expected utilities of acts *A* and *B* differ from the expected utilities of acts *C* and *D*, respectively, only by a constant amount ($.89^{*}u(\$1\text{M})$). The preference orderings should therefore coincide (that is, *A* is preferred to *B* if and only if *C* is preferred to *D*). More generally, any time two pairs of acts differ only by a common consequence *x* that obtains with probability *p*, the expected utilities of these pairs of acts differ by a constant, $p^{*}u(x)$. The preference ordering should therefore not be affected by the nature of this consequence (the value of *x*) or by its perceived likelihood (the value of *p*). Thus, when we consider the Ellsberg problem (see Table 14.3), we note that the preference between *G* and *H* should be the same as the preference between *E* and *F*, respectively, because the expected utilities of *G* and *H* differ from the expected utilities of *E* and *F*, respectively, by a constant amount, ($p(\text{blue})^{*}u(\$100)$).

The violations of the sure-thing principle discovered by Allais and Ellsberg both cast grave doubt on the descriptive validity of the classical model. However, the psychological intuitions underlying these violations are quite distinct. In Problem 1, people typically explain the apparent inconsistency as a preference for certainty: the difference between a 100 percent chance and a 99 percent chance of receiving a very large prize (*A* versus *B*) looms much larger than does the difference between an 11 percent chance and a 10 percent chance of receiving a very large prize (*C* versus *D*). The pattern of preferences exhibited for Problem 2, in contrast, seems to reflect a preference for known probabilities over unknown probabilities (or, more generally, a preference for knowledge over ignorance). In this case, *E* and *H* afford the decision maker precise probabilities, whereas *F* and *G* present the decision maker with vague probabilities.

The Allais problem suggests that people do not weight the utility of consequences by their respective probabilities as in the classical theory; the Ellsberg problem suggests that decision makers prefer more precise knowledge of probabilities. Both problems draw attention to deficiencies of the classical model in controlled environments where consequences are contingent on games of chance, such as a lottery or a drawing from an urn. Most real-world decisions, however, require decision makers to assess the probabilities of potential consequences themselves, with some degree of imprecision or vagueness. An important challenge to behavioral decision theorists over the past few decades has been to develop a more descriptively valid account of decision making that applies not only to games of chance but also to natural events, such as tomorrow's weather or the outcome of an election.

Our purpose in this chapter is to review a descriptive model of decision making under uncertainty. For simplicity, we will confine most of our discussion to acts entailing a single positive consequence (for example, receive \$100 if the home team wins and nothing otherwise). In Section 2, we take the Allais paradox as a point of departure and develop a psychological model of risky decision making that accommodates the preference for certainty. We extend these insights from situations where probabilities are provided to situations where decision makers must judge probabilities for themselves, and we develop a model that incorporates recent behavioral research on likelihood judgment. In Section 3, we take the Ellsberg paradox as a point of departure and describe a theoretical perspective that accommodates the preference for known probabilities. We extend this analysis to situations where consequences depend on natural events, and then modify the model developed in Section 2 to accommodate these new insights. Finally, in Section 4, we bring these strands together into a more unified account that distinguishes the role of beliefs, values and preferences in decision under uncertainty.

14.2 THE PREFERENCE FOR CERTAINTY: FROM ALLAIS TO THE TWO-STAGE MODEL

As we have observed, the Allais paradox violates the classical model of decision under uncertainty that weights utilities of consequences by their respective probabilities of occurrence. Moreover, numerous studies have shown that people often violate the principle of risk aversion that underlies much economic analysis. Table 14.4 illustrates a common pattern of risk aversion and risk seeking exhibited by participants in the studies of Tversky and Kahneman (1992). Let $C(x, p)$ be the "certainty equivalent" of the prospect (x, p) that offers to pay $\$x$ with probability p (that is, the sure payment that is deemed equally attractive to the prospect). The upper left-hand entry in the table shows that the median participant is indifferent between receiving \$14 for sure and a 5 percent chance of receiving \$100. Because the expected value of the prospect is only \$5, this observation reflects risk seeking.

Table 14.4 reveals a fourfold pattern of risk attitudes: risk seeking for low-probability gains and high-probability losses, coupled with risk aversion for high-probability gains and low-probability losses. Choices consistent with this pattern have been observed in several studies (e.g., Fishburn & Kochenberger, 1979; Kahneman & Tversky, 1979; Hershey & Schoemaker, 1980; Payne, Laughhunn & Crum, 1981). Risk seeking for low-probability gains may contribute to the attraction of gambling, whereas risk aversion for low-probability losses may contribute to the attraction of insurance. Risk aversion for high-probability gains may contribute to the preference for certainty in the Allais problem above (option *A* over option *B*), whereas risk seeking for high-probability losses is consistent with the common tendency to undertake risk to avoid facing a sure loss.

Table 14.4 The fourfold pattern of risk attitudes (adapted from Tversky & Kahneman, 1992). $C(x, p)$ is the median certainty equivalent of the prospect that pays $\$x$ with probability p

	Gain	Loss
Low probability	$C(\$100, .05) = \$14$ *risk seeking*	$C(-\$100, .05) = -\$8$ *risk aversion*
High probability	$C(\$100, .95) = \$78$ *risk aversion*	$C(-\$100, .95) = -\$84$ *risk seeking*

14.2.1 Prospect Theory's Weighting Function

The Allais paradox (Problem 1) cannot be explained by the shape of the utility function for money because options *A* and *B* differ from options *C* and *D* by a common consequence. Likewise, the fourfold pattern of risk attitudes (Table 14.4) cannot be explained by a utility function with both concave and convex regions (Friedman & Savage, 1948; Markowitz, 1952) because this pattern is observed over a wide range of payoffs (that is, a wide range of utilities). Instead, these patterns suggest a nonlinear transformation of the probability scale (cf. Preston & Baratta, 1948; Edwards, 1962), as advanced in prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992; other models with nonadditive probabilities include Quiggin, 1982; Gilboa, 1987; Schmeidler, 1989; Luce & Fishburn, 1992). According to prospect theory, the value V of a simple prospect that pays \$$x$ with probability p (and pays nothing with probability $1 - p$) is given by:

$$V(x, p) = v(x)w(p), \tag{2}$$

where v measures the subjective value of the consequence x, and w measures the impact of probability p on the attractiveness of the prospect. The value function, v, is a function of gains and losses relative to some reference point (usually the status quo), with $v(0) = 0$. The values of w are called decision weights; they are normalized so that $w(0) = 0$ and $w(1) = 1$. We pause to emphasize that w need not be interpreted as a measure of degree of belief—a person may believe that the probability of a fair coin landing heads is one-half but afford this event a weight of less than one-half in the evaluation of a prospect.

How might one measure the decision weight, $w(p)$? The simplest method is to elicit a person's certainty equivalent for a prospect that pays a fixed prize with probability p. For instance, suppose that Ann indicates that she is indifferent between receiving \$35 for sure, or receiving \$100 if a fair coin lands heads (and nothing if it lands tails). According to prospect theory (Equation (2)):

$$v(35) = v(100)w(.5)$$

so that

$$w(.5) = v(35)/v(100).$$

Now, to simplify our analysis, let us suppose that Ann's value function is linear,[6] so that $v(x) = x$. In this case:

$$w(.5) = .35.$$

Hence, in this example, a .5 probability receives a weight of .35 in the evaluation of the prospect, and Ann's risk aversion would be attributed not to the shape of the value function (as in expected utility theory—see Figure 14.1), but rather to the underweighting of a .5 probability.

According to prospect theory, the shapes of both the value function $v(.)$ and weighting function $w(.)$ reflect psychophysics of diminishing sensitivity: marginal impact diminishes with distance from the reference point. For monetary outcomes, the status quo generally serves as the reference point distinguishing losses from gains, so that the function is concave for gains and convex for losses (see Figure 14.2a). Concavity for gains contributes to risk aversion for gains (as we saw in the analysis of the concave utility function in Figure 14.1), and convexity for losses contributes to risk seeking for losses. The prospect theory value

[6] A more typical individual, as we shall see, can be characterized instead by a concave value function for gains; for example, $v(x) = x^{\alpha}$, $0 < \alpha < 1$. For relatively small dollar amounts such as those presented in this example, however, $v(x) = x$ is a reasonable first-order approximation, so that risk attitudes are driven primarily by the weighting of probabilities. However, the studies reviewed later in this chapter do not rely on such a simplifying assumption.

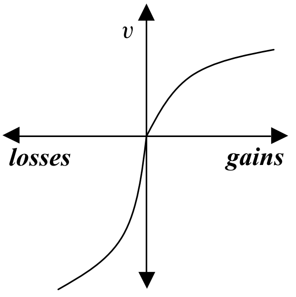

Figure 14.2a Value function, v, for monetary gains and losses

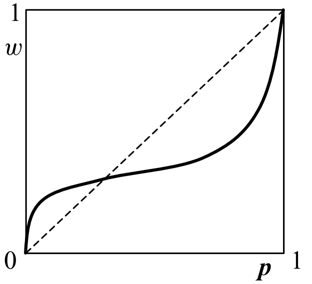

Figure 14.2b Weighting function, w, for chance events that obtain with probability p

function is also steeper for losses than gains. This gives rise to risk aversion for mixed (gain-loss) gambles, so that, for example, people typically reject a gamble that offers a .5 chance of gaining \$100 and a .5 chance of losing \$100. As noted earlier, we will confine most of the analysis in this chapter to potential gains, so we will postpone a further discussion of losses until the conclusion.

For probability, there are two natural reference points: impossibility and certainty. Hence, diminishing sensitivity implies an inverse-S shaped weighting function that is concave near zero and convex near one, as depicted in Figure 14.2b. It explains the fourfold pattern of risk attitudes (Table 14.4), because low probabilities are overweighted (leading to risk seeking for gains and risk aversion for losses), and high probabilities are underweighted (leading to risk aversion for gains and risk seeking for losses).[7] It also accounts for the Allais paradox (Problem 1), because $w(1) - w(.99) >> w(.11) - w(.10)$. That is, increasing the probability of winning a large prize from .99 to 1 has more impact on the decision maker than increasing the probability of winning from .10 to .11. This inverse-S shaped weighting function seems to be consistent with a range of empirical findings (see Camerer & Ho, 1994; Wu & Gonzalez, 1996, 1998; Abdellaoui, 2000; Wakker, 2000; for parameterization of the weighting function, see Prelec, 1998; Gonzalez & Wu, 1999; for applications to decision analysis, see Bleichrodt & Pinto, 2000; Bleichrodt, Pinto & Wakker, 2001).

[7] As we stated in the previous paragraph, a value function that is concave for gains and convex for losses implies risk aversion for gains and risk seeking for losses. This pattern is reinforced by a weighting function that underweights moderate to large probabilities, but it is reversed by a weighting function that overweights low probabilities.

14.2.2 From Risk to Uncertainty: Measuring Diminished Sensitivity

The inverse-S shaped weighting function provides a parsimonious account of decision making in situations where outcome probabilities are known with precision by the decision maker. However, most decisions (as we have already observed) require the decision maker to assess probabilities herself, with some degree of imprecision or vagueness. Following Knight (1921), theorists distinguish between decisions under *risk*, where probabilities are known, and decisions under *uncertainty*, where probabilities are not known. The question arises of how to extend the analysis of decision weights from risk to uncertainty, because under uncertainty we can no longer describe decision weights as a transformation of the probability scale.

One approach to solving this problem is to formalize the notion of diminishing sensitivity for risk, and then extend this definition to uncertainty. Diminishing sensitivity means that the weight of an event decreases with distance from the natural boundaries of zero and one. Let p, q and r be numbers such that $0 < p, q, r < 1, p + q + r < 1$. Diminishing sensitivity near zero can be expressed as:

$w(p) - w(0) \geq w(p + q) - w(q) \geq w(p + q + r) - w(q + r)$, and so forth.

That is, adding probability p to an impossibility of winning a prize has a greater impact than adding p to some intermediate probability q, and this, in turn, has a greater impact than adding p to some larger probability $q + r$, and so forth. In general, this "diminishing" sensitivity is most pronounced near the boundaries (that is, the pattern expressed by the leftmost inequality above is more robust than the pattern expressed by subsequent inequalities). Hence, we will focus our attention on the "diminished" sensitivity to intermediate changes in outcome probabilities near zero. Noting that $w(0) = 0$, diminished sensitivity near zero can be expressed as:

$$w(p) \geq w(p + q) - w(q). \tag{3a}$$

Similarly, noting that $w(1) = 1$, diminished sensitivity near one can be expressed as:

$$1 - w(1 - p) \geq w(1 - q) - w(1 - q - p). \tag{3b}$$

To illustrate, suppose $p = q = .1$. The lower left-hand corner of Figure 14.3 illustrates "lower subadditivity". The impact of .1 is greater when added to 0 than when added to .1 (the length of segment *A* is greater than the length of segment *B*). For instance, consider a lottery with 10 tickets so that each ticket has a .1 chance of winning a fixed prize. Most people would pay more for a ticket if they did not have one (improving the probability of winning from 0 to .1) than they would pay for a second ticket if they already had one (improving the probability of winning from .1 to .2).[8] The upper right-hand corner in Figure 14.3 illustrates "upper subadditivity". The impact of .1 is greater when subtracted from 1 than when subtracted from .9 (the length of segment *C* is greater than the length of segment *D*). For instance, most people would pay more for the tenth ticket if they already had nine (improving the probability of winning from .9 to 1) than they would pay for a ninth ticket if they already had eight (improving the probability of winning from .8 to .9).[9]

This pattern of diminished sensitivity can be readily extended from risk to uncertainty. Again, let *S* be a set whose elements are interpreted as states of the world. Subsets of *S* are

[8] Such a pattern could be accommodated in expected utility theory only through a convex utility function for money.
[9] For an empirical demonstration similar to the lottery anecdote used here, see Gonzalez and Wu (1999).

called "events". Thus, S corresponds to the certain event, Ø is the null event (that is, the impossible event), and $S - A$ is the complement of event A. A weighting function W (on S) is a mapping that assigns to each event in S a number between 0 and 1 such that $W(\emptyset) = 0$, $W(S) = 1$, and $W(A) \geq W(B)$ if A includes B. Note that the weighting function for uncertain events, W, should be distinguished from the weighting function for risky (chance) events, w. Thus, for uncertainty, we can rewrite Equation (2) so that the value of prospect (x, A) that offers \$$x$ if event A obtains and nothing otherwise is given by:

$$V(x, A) = v(x)W(A).$$

Equation (3a) can now be generalized as *lower subadditivity:*

$$W(A) \geq W(A \cup B) - W(B), \tag{4a}$$

provided A and B are disjoint (that is, mutually exclusive), and $W(A \cup B)$ is bounded away from one.[10] This inequality is a formal expression of the *possibility effect*: the impact of event A is greater when it is added to the null event than when it is added to some non-null event B. Equation (3b) can be generalized as *upper subadditivity*:

$$1 - W(S - A) \geq W(S - B) - W(S - A \cup B), \tag{4b}$$

provided $W(S - A \cup B)$ is bounded away from zero. Upper subadditivity[11] is a formal expression of the *certainty effect*: the impact of event A is greater when it is subtracted from the certain event S than when it is subtracted from some uncertain event $S - B$. Note that upper subadditivity can be expressed as lower subadditivity of the dual function, $W'(A) \equiv 1 - W(S - A)$. That is, upper subadditivity is the same as lower subadditivity where we transform both scales by subtracting events from certainty and decision weights from one (reversing both axes, as can be seen by viewing Figure 14.3 upside down).

Why the terms "lower subadditivity" and "upper subadditivity"? "Lower" and "upper" distinguish diminished sensitivity near zero (the lower end of the scale) from diminished sensitivity near one (the upper end of the scale), respectively. "Subadditivity" refers to the implication revealed when we rearrange terms of Equations (4a) and (4b):

$$W(A \cup B) \leq W(A) + W(B) \tag{i}$$

and

$$W'(A \cup B) \leq W'(A) + W'(B). \tag{ii}$$

Thus, when disjoint events are concatenated (added together) they receive less weight than when they are weighted separately and summed—W is a *sub-additive* function of events. The weighting function satisfies *bounded subadditivity*, or subadditivity (SA) for short, if it satisfies both (4a) and (4b).[12]

---

[10] The boundary conditions are needed to ensure that we always compare an interval that includes an endpoint (zero or one) to an interval that does not include an endpoint. See Tversky and Wakker (1995) for a more formal discussion.

[11] Note that if we define $B' = S - A \cup B$, then upper subadditivity can be expressed as $1 - W(S - A) \geq W(A \cup B') - W(B')$. Upper subadditivity has been previously presented in this form (Tversky & Fox, 1995; Tversky & Wakker, 1995; Fox & Tversky, 1998).

[12] For a more formal treatment of bounded subadditivity, see Tversky and Wakker (1995). For a more thorough account of diminishing sensitivity under risk that explores concavity near zero, convexity near one and diminishing marginal concavity throughout the scale, see Wu and Gonzalez (1996); for extensions to uncertainty, see Wu and Gonzalez (1999b).

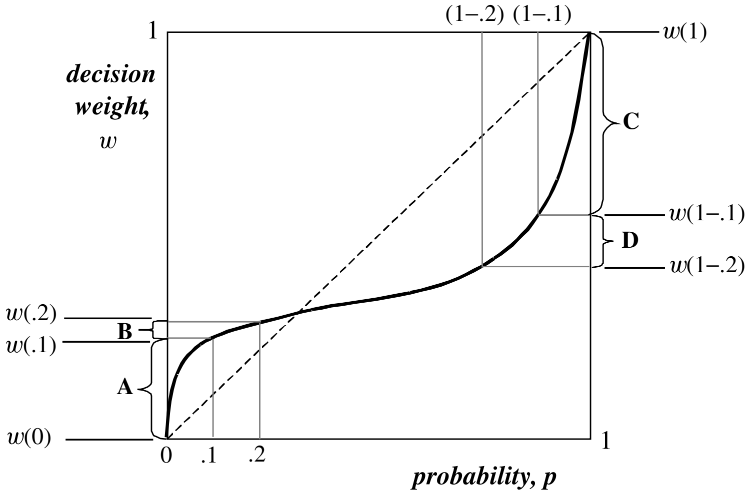

Figure 14.3 Visual illustration of bounded subadditivity. The lower left-hand corner of the figure illustrates lower subadditivity. The upper right-hand corner illustrates upper subadditivity

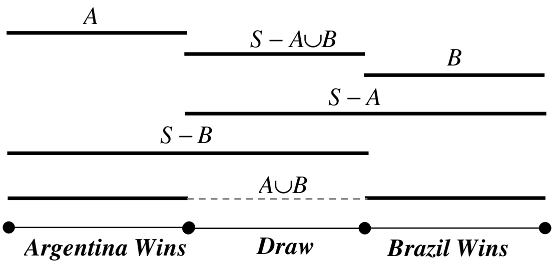

Figure 14.4 An event space for prospects defined by the result of a soccer match between Argentina and Brazil. Each row denotes a target event that defines a prospect

To illustrate bounded subadditivity more concretely, consider a soccer match between Argentina and Brazil. We can partition the state space into three elementary events (see Figure 14.4): Argentina wins (A), Brazil wins (B), or there is a draw ($S - A \cup B$). Additionally, this partition defines three compound events: Argentina fails to win ($S - A$), Brazil fails to win ($S - B$), and there is a decisive game (that is, either Argentina or Brazil wins, $A \cup B$). Now suppose that we ask a soccer fan to price prospects that would pay \$100 if each of these target events obtains. For instance, we ask the soccer fan what sure amount of money, C_A, she finds equally attractive to the prospect (\$100, A) that offers \$100 if Argentina wins (and nothing otherwise). Let us suppose further, for simplicity, that this individual values money according to a linear value function, so that $v(x) = x$. In this case, $W(A) = C_A/100$.

Suppose our soccer fan prices bets on Argentina winning, Brazil winning, and a decisive match at \$50, \$40 and \$80, respectively. In this case, we get $W(A) = .5$, $W(B) = .4$, and $W(A \cup B) = .8$, so that lower subadditivity (Equation (4a)) is satisfied because $.5 > .8 - .4$. Suppose further that our soccer fan prices bets on Argentina failing to win, Brazil failing to win, and a draw at \$40, \$50 and \$10, respectively. In this case, we get $W(S - A) = .4$,

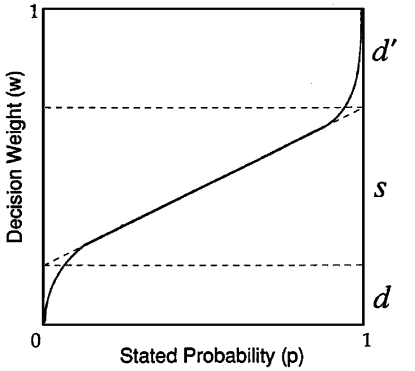

Figure 14.5 A weighting function that is linear except near the endpoints (d = "lower" intercept of the weighting function; d' = "upper" intercept of the weighting function; s = slope. Reproduced from Tversky & Fox, 1995)

$W(S - B) = .5$, and $W(S - A \cup B) = .1$, so that upper subadditivity (Equation (4b)) is satisfied because $1 - .4 > .5 - .1$.

We can determine the degree of subadditivity by assessing the magnitude of the discrepancy between terms on either side of inequalities (*i*) and (*ii*) above:

$$D(A, B) \equiv W(A) + W(B) - W(A \cup B)$$
$$D'(A, B) \equiv W'(A) + W'(B) - W'(A \cup B).$$

This metric provides a measure of the departures from additivity of the weighting function around impossibility (D) and certainty (D'). These measures are particularly useful because they do not require specification of objective probability. We can thus compare the degree of subadditivity between risk and uncertainty. More generally, we can compare the degree of subadditivity between different domains or *sources* of uncertainty, where a source of uncertainty is interpreted as a family of events that are traced to a common causal system, such as the roll of a die, the winner of an election or the final score of a particular soccer match.[13]

Suppose that an experimenter measures decision weights of several events so that she has multiple tests of upper and lower subadditivity for each individual. To obtain summary measures of subadditivity, let d and d', respectively, be the mean values of D and D' for a given respondent and source of uncertainty. To see how one might interpret the values of d and d', assume a weighting function that is approximately linear except near the endpoints (see Figure 14.5). It is easy to verify that within the linear portion of the graph, D and D' do not depend on A and B, and the values of d and d' correspond to the 0- and 1-intercepts of such a weighting function. Thus, d measures the magnitude of the impossibility "gap" and d' measures the magnitude of the certainty "gap". Moreover, we can define a global index of sensitivity, $s = 1 - d - d'$ that measures the slope of the weighting function—that is, a person's sensitivity to changes in probability. Prospect theory assumes that $d \geq 0$, $d' \geq 0$, and $s \leq 1$, whereas expected utility theory assumes $d = d' = 0$, and $s = 1$. An extreme instance in which $s = 0$ (and also $d > 0$, $d' > 0$) would characterize a three-valued logic in which a person distinguishes only impossibility, possibility and certainty.

[13] In decision under risk, we interpret uncertainty as generated by a standard random device; although probabilities could be realized through various devices, we do not distinguish between them and instead treat risk as a single source.

Tversky and Fox (1995) tested bounded subadditivity in a series of studies using risky prospects (for example, receive \$150 with probability .2) and uncertain prospects with outcomes that depended on future temperature in various cities, future movement of the stock market, and the result of upcoming sporting events (for example, "receive \$150 if the Buffalo Bills win the Super Bowl"). These authors estimated certainty equivalents by asking participants to choose between each prospect and a series of sure payments. For example, if a participant favored \$35 for sure over a prospect that offered \$150 with probability .2, and that participant also favored the prospect over receiving \$30 for sure, then the certainty equivalent for the prospect was estimated to be \$32.50 (midway between \$30 and \$35). The authors then estimated decision weights as $W(A) = v(C_A)/v(150)$, where the value function was estimated using data from another study. The results were very consistent: bounded subadditivity was pronounced for risk and all sources of uncertainty ($d > 0$, $d' > 0$ for a significant majority of participants). In addition, Tversky and Fox found that subadditivity was more pronounced for uncertainty than for risk. That is, values of d and d' were larger for uncertain prospects than for risky prospects.

Note that Figure 14.5, like Figures 14.2b and 14.3, is drawn so that the weighting function crosses the identity line below .5, and $d' > d$. This is meant to reflect the assumption under prospect theory that decision weights for complementary events generally sum to less than one, $W(A) + W(S - A) \leq 1$, or equivalently, $W(A) \leq W(S) - W(S - A)$. This property, called "subcertainty" (Kahneman & Tversky, 1979), accords with the data of Tversky and Fox (1995) and can be interpreted as evidence of more pronounced upper subadditivity than lower subadditivity; that is, the certainty effect is more pronounced than the possibility effect.

Figure 14.6 summarizes results from Tversky and Fox (1995), plotting the sensitivity measure s for risk versus uncertainty for all participants in their studies. Three patterns are worth noting. First, $s < 1$ for all participants under uncertainty (mean $s = .53$) and for all but two participants under risk (mean $s = .74$). Second, the value of s is smaller under uncertainty than under risk for 94 of 111 participants (that is, the large majority of points lie below the identity line). Third, there is a significant correlation between sensitivity to risk and sensitivity to uncertainty ($r = .37, p < .01$).

14.2.3 The Two-Stage Model

The observation that subadditivity is more pronounced under uncertainty than risk accords with the natural intuition that people could be less sensitive to changes in the target event when they do not have objective probabilities at their disposal but must instead make a vague assessment of likelihood. Fox and Tversky (1995) speculated that increased subadditivity under uncertainty might be directly attributable to subadditivity of judged probabilities. To test this conjecture, they asked participants in all of their studies to judge the probabilities of all target events. These researchers found that judged probability, $P(.)$, exhibited significant bounded subadditivity.[14] That is, if D_P and D'_P measure the degree of lower and upper subadditivity, respectively, of judged probabilities for disjoint events A and B, we get:

$$D_P(A, B) \equiv P(A) + P(B) - P(A \cup B) > 0$$

and

[14] Note that we distinguish judged probability, $P(.)$, from objective probability, p.

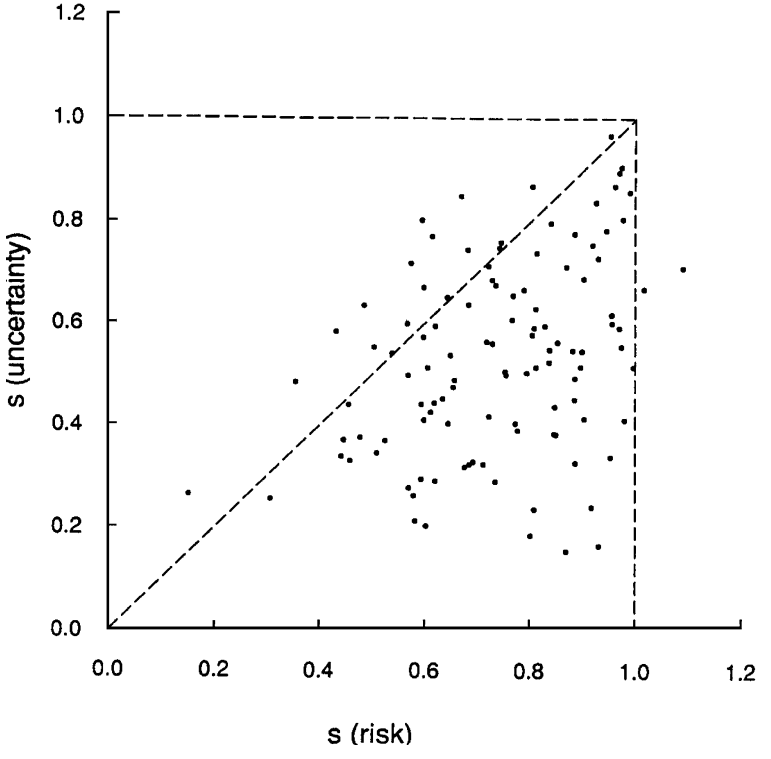

Figure 14.6 A plot of the joint distribution of the sensitivity measure *s* for risk and uncertainty for participants in the studies of Tversky and Fox (1995). Reproduced from Tversky and Fox (1995)

$$D'_P(A, B) \equiv P'(A) + P'(B) - P'(A \cup B) > 0$$

(where $P'(A) = 1 - P(S - A)$), for a significant majority of tests for all sources of uncertainty. Interestingly, the degree of subadditivity observed for direct judgments of probability was less than the degree of subadditivity observed for decision weights that were inferred from choices. This observation motivated Fox and Tversky (1995) to propose a two-stage model in which decision makers first judge the probability P of the target event A, and then transform this judged probability by the risky weighting function w. Thus, according to the two-stage model,

$$W(A) = w[P(A)]. \tag{5}$$

Indeed, when Fox and Tversky (1995) plotted the uncertain decision weight, $W(A)$, for each target event against the corresponding judged probability, $P(A)$, these plots closely resembled the plot of the risky weighting function, $w(p)$, against objective probability p for the same group of participants. The notion that risky decision weights are a constituent of uncertain decision weights may also explain the aforementioned finding of a significant positive correlation between sensitivity to risk and sensitivity to uncertainty (see Figure 14.6).

Further evidence for the two-stage model was obtained in a study of professional options traders who were surveyed on the floors of the Pacific Stock Exchange and Chicago Board Options Exchange by Fox, Rogers and Tversky (1996). Options traders are unique in that they are schooled in the calculus of chance and make daily decisions under uncertainty on which their jobs depend. Unlike participants in most studies, the majority of the options traders priced risky prospects by their expected value. This pattern is consistent with both a linear value function and linear risky weighting function.[15] However, when participants were asked to price uncertain prospects contingent on future stock prices (for example, "receive $150 if Microsoft stock closes below $88 per share two weeks from today"), their decision weights exhibited pronounced subadditivity. Furthermore, when these same participants were asked to judge the probability of each target event, they exhibited roughly the same degree of subadditivity as they had exhibited for decision weights. Thus, when decision weights were plotted against judged probabilities, points fell roughly along the identity line (that is, $W(A) = P(A)$). This pattern is consistent with the two-stage model, in which options traders first judge the probability of each target event—subadditively—and then weight the $150 prize by this judged probability.

Fox and Tversky (1998) elaborated the two-stage model (Equation (5)) and tested some of its implications. In this theory, both the uncertain weighting function, $W(.)$, and the risky weighting function, $w(.)$, are assumed to conform to prospect theory (that is, satisfy bounded subadditivity). In addition, judged probability, $P(.)$, is assumed to conform to support theory (Tversky & Koehler, 1994; Rottenstreich & Tversky, 1997), a descriptive model of judgment under uncertainty. To demonstrate how support theory accommodates subadditivity of judged probability and to highlight its novel implications for the two-stage model, we describe the key features of support theory in the section that follows.

Support Theory

There is abundant evidence from prior research that judged probabilities do not conform to the laws of chance (e.g., Kahneman, Slovic & Tversky, 1982). In particular, alternative descriptions of the same event can give rise to systematically different probability judgments (e.g., Fischhoff, Slovic & Lichtenstein, 1978), more inclusive events are sometimes judged to be less likely than less inclusive events (Tversky & Kahneman, 1983), and the judged probability of an event is typically less than the sum of judged probabilities of constituent events that are evaluated separately (e.g., Teigen, 1974). To accommodate such patterns, support theory assumes that judged probability is not attached to events, as in other theories, but rather to descriptions of events, called "hypotheses". Thus, two different descriptions of the same event may be assigned distinct probabilities. Support theory assumes that each hypothesis A has a non-negative support value $s(A)$ corresponding to the strength of evidence for that hypothesis. Support is assumed to be generated through heuristic processing of information or through explicit reasoning or computation (see also Sloman et al, 2003). The judged probability $P(A, \bar{A})$ that hypothesis A, rather than its complement $\bar{A}$, obtains is given by:

[15] Fox, Rogers and Tversky (1996) claimed that this pattern of results would be observed under cumulative prospect theory *if and only if* the value function and weighting function were both linear (p. 8, lines 15–18). Fox and Wakker (2000) observe that this assertion is not technically correct given the method by which the value function was elicited in that study, but that the conclusion is pragmatically reasonable and that the qualitative results reported in that paper are robust over a wide range of variations in the value function. For a method of measuring the shape of the value function that does not assume additive subjective probabilities, see Wakker and Deneffe (1996). A nonparametric algorithm for simultaneously estimating subjective value and decision weights from certainty equivalents of risky prospects was advanced by Gonzalez and Wu (1999).

$$P(A, \bar{A}) = s(A)/[s(A) + s(\bar{A})]. \tag{6}$$

This equation suggests that judged probability can be interpreted as the balance of evidence for a focal hypothesis against its alternative. Hence, if the support for a hypothesis (for example, "rain tomorrow") and its complement (for example, "no rain tomorrow") are equal, the judged probability is one-half (that is, P(rain, no rain) = .5). As the support for the focal hypothesis increases relative to support for the alternative hypothesis, judged probability approaches one. Likewise, as support for the alternative hypothesis increases relative to support for the focal hypothesis, judged probability approaches zero.

The theory further assumes that (i) unpacking a hypothesis A (for example, "the winner of the next US presidential election will not be a Democrat") into an explicit disjunction of constituent hypotheses $A_1 \vee A_2$ (for example, "the winner of the next US presidential election will be a Republican or an independent candidate") generally increases support, and (ii) separate evaluation of the constituent hypotheses (for example, "the winner of the next US presidential election will be a Republican"; "the winner of the next US presidential election will be an independent candidate") generally gives rise to still higher total support. More formally,

$$s(A) \leq s(A_1 \vee A_2) \leq s(A_1) + s(A_2), \tag{7}$$

provided (A_1, A_2) is recognized as a partition (that is, exclusive and exhaustive constituents) of A.

The first set of consequences of support theory concerns the additivity of judged probabilities. Equation (6) implies *binary complementarity*: $P(A) + P(\bar{A}) = 1$. That is, the judged probability of A and its complement sum to unity.[16] For instance, the judged probability that the winner of the next election will be a "Democrat" plus the judged probability that the winner of the next election will "not be a Democrat" should sum to one. For finer partitions, however, Equations (6) and (7) imply *subadditivity:* $P(A) \leq P(A_1) + P(A_2)$. That is, the probability of hypothesis A is less than or equal to the sum of probabilities of its disjoint components (note that this also implies that the judged probabilities of $n > 2$ exhaustive and exclusive hypotheses generally sum to more than one). For instance, the judged probability that the winner of the next election will "not be a Democrat" is less than or equal to the judged probability of "Republican" plus the judged probability of "an independent candidate". Such patterns have been confirmed in several studies reviewed by Tversky and Koehler (1994). Subadditivity and binary complementarity have also been documented in studies of physicians (Redelmeier et al., 1995), lawyers (Fox & Birke, 2002) and options traders (Fox, Rogers & Tversky, 1996); and subadditivity has been observed in published odds of bookmakers (Ayton, 1997). A within-subject test that traces the relationship between raw expression of support and judged probabilities is reported by Fox (1999; see also Koehler, 1996). For a demonstration of subadditivity in a classification learning task, see Koehler (2000). For exceptions to binary complementarity, see Brenner and Rottenstreich (1999), Macchi, Osherson and Krantz (1999) and Idson et al. (2001).

[16] In the studies that we will review here, participants are asked to evaluate a target hypothesis A (for example, judge the probability that "the Lakers win the NBA championship"), so that the alternative hypothesis is not explicitly identified (that is, participants are not asked to judge the probability that "the Lakers win rather than fail to win the NBA championship"). However, we assume that when evaluating the likelihood of a target event A, decision makers consider its simple negation $\bar{A}$ as the default alternative hypothesis. Moreover, to simplify our exposition, we abbreviate the expression $P(A, \bar{A})$ in our notation as $P(A)$, taking the negation of the focal hypotheses, $\bar{A}$, as implicit. Also, to simplify we use letters A, B, etc., henceforth to refer to hypotheses (that is, descriptions of events) rather than subsets of a state space.

A second important consequence of support theory is *implicit subadditivity*: $P(A) \leq P(A_1 \vee A_2)$. That is, the judged probability of a hypothesis generally increases when unpacked into an explicit disjunction of constituent hypotheses. For instance, the judged probability that the winner of the next election will "not be a Democrat" is less than or equal to the judged probability that the winner will be a "Republican or independent candidate". This pattern has been demonstrated empirically by Rottenstreich and Tversky (1997) and replicated by others (Redelmeier et al., 1995; Fox & Tversky, 1998; Fox & Birke, 2002; for counterexamples, see Sloman et al., 2003).

Because the two-stage model incorporates judged probabilities from support theory, it is a significant departure from other models of decision under uncertainty in two respects. First, subadditivity of the uncertain weighting function, $W(.)$, is partially attributed to subadditivity of judged probability. Traditional economic models of decision under uncertainty infer decision weights from observed choices and therefore cannot distinguish between belief and preference components of decision weights. Second, $W(.)$ allows different decision weights for distinct descriptions of a target event: $W(A) \leq W(A_1 \vee A_2)$. Thus, different descriptions of options can give rise to different preferences, in violation of the normative assumption of *description invariance* (see Tversky & Kahneman, 1986; Tversky, 1996).

The Partition Inequality

One of the most striking differences between the classical model of decision under uncertainty and the two-stage model can be observed directly from certainty equivalents for uncertain prospects. Suppose $(A_1, A_2, \ldots, A_n)$ is recognized as a partition of hypothesis A, and that $C(x, A)$ is the certainty equivalent of the prospect that pays $\$x$ if hypothesis A obtains, and nothing otherwise. Expected utility theory assumes that the certainty equivalent of an uncertain prospect is not affected by the particular way in which the target event is described, so that:

$$C(x, A) = C(x, A_1 \vee \ldots \vee A_n) \qquad (i)$$

for all real x, $n > 1$. For instance, a person's certainty equivalent for the prospect that pays $100 if "there is measurable precipitation next April 1 in Chicago" should be the same as the certainty equivalent for the prospect that pays $100 if "there is measurable rain or sleet or snow or hail next April 1 in Chicago".

Moreover, expected utility theory with risk aversion implies:

$$C(x, A) \geq C(x, A_1) + \cdots + C(x, A_n). \qquad (ii)$$

That is, the certainty equivalent for a prospect is at least as large as the sum of certainty equivalents for subprospects that are evaluated separately (for a derivation, see Fox & Tversky, 1998, p. 882, Footnote 6). To illustrate, note that for a risk-averse individual the certainty equivalent of a prospect that pays $100 if a fair coin lands heads will be less than its expected value of $50. Likewise, for this same individual the certainty equivalent of a prospect that pays $100 if a fair coin lands tails will be less than $50. Hence the aggregate price of the subprospects, evaluated separately by this risk-averse individual, will be less than the price of the prospect that pays $100 for sure.

Taken together, (*i*) and (*ii*) above imply the following *partition inequality*:[17]

$$C(x, A) = C(x, A_1 \vee \ldots \vee A_n) \geq C(x, A_1) + \ldots + C(x, A_n). \quad (8)$$

If decision makers act in accordance with the two-stage model (Equation (5)), then the partition inequality will not generally hold. In particular, situations can arise where:

$$C(x, A) < C(x, A_1 \vee \ldots \vee A_n) < C(x, A_1) + \ldots + C(x, A_n).$$

The equality (*i*) in Equation (8) is especially likely to fail when the explicit disjunction reminds people of possibilities that they may have overlooked or makes compelling possibilities more salient. The inequality (*ii*) in Equation (8) is especially likely to fail when the target event is partitioned into many pieces so that subadditivity of judged probability and subadditivity of the risky weighting function are more pronounced than concavity of the value function.

14.2.4 Tests of the Two-Stage Model

We now present evidence from several previous studies that test the two-stage model against the classical model (expected utility with risk aversion). We begin with tests of the partition inequality, and then proceed to more direct tests of the relative fit of these models.

Tests of the Partition Inequality

To date, only partial tests of the entire partition inequality (Equation (8)) have been published. In this section, we begin by presenting evidence from previous studies for violations of description invariance (part (*i*)) and risk aversion (part (*ii*)), and then present evidence from new studies that simultaneously test the entire pattern in Equation (8) within-subject.

Violations of Description Invariance

There have been only a few documented tests of unpacking effects (implicit subadditivity) against the description invariance assumption of expected utility theory (part (*i*) of Equation (8)). Johnson et al. (1993) reported that consumers were willing to pay more for an insurance policy that covered hospitalization due to “any disease or accident” ($47.12) than hospitalization for “any reason” ($41.53). Similarly, Fox and Tversky (1998) found that participants valued the prospect offering $75 if “Chicago or New York or Indiana or Orlando” wins the National Basketball Association (NBA) playoffs more highly than the prospect offering $75 if an “Eastern Conference” team wins the NBA playoffs, despite the fact that the teams listed in the former prospect were only a (proper) subset of the teams comprising the Eastern Conference.

In a more extensive investigation, Wu and Gonzalez (1999a) observed that participants favored “unpacked” prospects over “packed” prospects, even though the latter dominated the former. For instance, most participants said they would rather receive a sure payment

[17] This version of the partition inequality extends Equation (4) from Fox and Tversky (1998).

Table 14.5 Subadditivity and violations of the partition inequality

				Certainty equivalents		Judged probability	
Study	N	Source of uncertainty	Atoms	ΣC	$\Sigma C(A_i) > C(S)$	$P(A) + P(S - A)$	ΣP
Tversky & Fox (1995)							
NBA fans	27	a. NBA playoffs	6	1.40	93%	0.99	1.40
		b. San Francisco temperature	6	1.27	77%	0.98	1.47
NFL fans	40	c. Super Bowl	6	1.31	78%	1.01	1.48
		d. Dow Jones	6	1.16	65%	0.99	1.25
Stanford students	45	e. San Francisco temperature	6	1.98	88%	1.03	2.16
		f. Beijing temperature	6	1.75	82%	1.01	1.88
Fox et al. (1996)							
Options traders (Chicago)	32	g. Microsoft	4	1.53	89%	1.00	1.40
		h. GE	4	1.50	89%	0.96	1.43
Options traders (SF)	28	i. IBM	4	1.47	82%	1.00	1.27
		j. Gannett Co.	4	1.13	64%	0.99	1.20
Fox & Tversky (1998)							
NBA fans	50	k. NBA playoffs	8	2.08	82%	1.02	2.40
Stanford students	82	l. Economic indicators	4	1.14	41%	0.98	1.14
Median				1.44	82%	1.00	1.42

Note: Adapted from Fox and Tversky (1998, Table 8). The first three columns identify the participant population, sample size and sources of uncertainty. The fourth column lists the number of elementary events (*atoms*) in the finest partition of the state space (S) available in each study. The fifth and sixth columns present the median sum of normalized certainty equivalents for an *n*-fold partition of S, and the proportion of participants who reported certainty equivalents that summed to more than the amount of the prize. The last two columns present the median sum of judged probabilities for binary partitions of S and an *n*-fold partition of S.

of \$150 than a prospect that offered \$*240 if the Bulls win more than 65 games (and nothing otherwise)*. However, when the prospect was modified so that it offered \$*240 if the Chicago Bulls win between 65 and 69 games and* \$*220 if the Bulls win more than 70 games (and nothing otherwise)*, most participants instead favored the prospect over a sure payment of \$150. Apparently, unpacking the target prospect increased its attractiveness even though the unpacked prospect is dominated by the packed prospect (note that the unpacked version offers \$20 less if the Bulls win more than 70 games). Wu and Gonzalez (1999a) obtained a similar pattern of results for prospects drawn from diverse domains such as election outcomes and future temperatures. Although, strictly speaking, Wu and Gonzalez's demonstrations do not entail a violation of description invariance, they do suggest that unpacking the target event into a separate description of constituent events (each with similar consequences) can increase the attractiveness of a prospect.

Violations of Risk Aversion

There have been a number of published studies testing (lower) subadditivity against expected utility with risk aversion (part (*ii*) of Equation (8)). Table 14.5 summarizes these results, showing a pattern that consistently favors the two-stage model. First, consider certainty equivalents. The column labeled ΣC presents the median sum of normalized certainty equivalents (that is, certainty equivalents divided by the amount of the relevant prize) for the finest partition of S available in each study, and the column labeled $\Sigma C(A_i) > C(S)$ presents the corresponding percentage of participants who violated part (*ii*) of the partition inequality for the finest partition of S. In accord with the two-stage model, the majority of participants in every study violated the partition inequality, and the sum of certainty equivalents was often substantially greater than the prize. For instance, in Fox and Tversky's (1998) sample of NBA fans, participants priced prospects that offered \$160 if a particular team would win the NBA playoffs (eight teams remained in contention at the time of the study). The median participant reported certainty equivalents for the eight teams that sum to 2.08 times the \$160 prize (that is, \$330). Moreover, 82 percent of respondents violated part (*ii*) of the partition inequality by reporting certainty equivalents that exceed \$160.

Next, consider probability judgment. The column labeled $P(A) + P(S{-}A)$ presents the median sum of judged probabilities for binary partitions of the state space. The column labeled ΣP presents the median sum of judged probabilities for the finest partition of S. The results conform to support theory: sums for binary partitions are close to one, whereas sums for finer partitions are consistently greater than one. For instance, in Fox and Tversky's (1998) study of NBA fans, the median sum of probabilities that the "Western Conference" would win the NBA championship and the "Eastern Conference" would win the NBA championship was 1.02. However, when these same participants separately judged the probabilities that each of the eight teams remaining would win the NBA championship, the median participant reported numbers that summed to 2.40.

Violations of the Partition Inequality

In a new set of studies, we (Fox & See, 2002) have documented violations of the entire partition inequality (Equation (8)), including tests of both description invariance (part (*i*)) and risk aversion (part (*ii*)). We replicated the methods of Fox and Tversky (1998), adding elicitation of explicit disjunctions (the middle term in Equation (8)) so that the entire pattern predicted by support theory and the two-stage model could be tested simultaneously and within-subject.

The data presented here are drawn from two studies. In the first study, Duke University undergraduates were asked to price prospects that offered \$160 depending on which team would win the Atlantic Coast Conference (ACC) men's college basketball tournament (a topic that Duke students follow closely), and then judge probabilities of all target events. The ACC consists of nine schools in the southeastern USA. Schools were categorized according to geographic subregion ("inside North Carolina" versus "outside North Carolina") and funding ("private school" versus "public school").[18] In the second study, Brown University students first learned the movement of economic indicators in a hypothetical economy (interest rates and unemployment, which could each move either "up" or "down" in a given period). Next they were asked to price prospects that offered \$160 depending on the direction in which indicators would move in the following period. Finally, participants were asked to judge probabilities of all target events. Participants in

[18] The results presented here from Fox and See (2002) are a subset of results from this study in which we relied on a hierarchical partition of the state space. Participants were also asked to price prospects contingent on the final score of an upcoming game (a dimensional partition) and prospects contingent on the conjunction of results in two upcoming games (a product partition). The results of these latter tasks were consistent with the two-stage model, but did not yield significant implicit subadditivity in either judgment or choice.

Table 14.6 Mean certainty equivalents for joint events, explicit disjunction, and sums

Study	N	Source of uncertainty		Implicit (A)		Explicit ($A_1 \vee A_2$)		Sum ($A_1 + A_2$)
Fox & See (2002)								
Duke basketball fans	51	Basketball	$P(.)$	0.56	**	0.65	**	0.81
			$C(.)$	0.53	**	0.61	**	0.76
Brown students	29	Economic indicators	$P(.)$	0.52	**	0.55	*	0.68
			$C(.)$	0.42	†	0.52	**	0.63

† $p < .10$; *$p < .05$; **$p < .005$.
The first three columns identify the participant population, sample size and sources of uncertainty. The fourth column indicates the relevant dependent variable (judged probability or normalized certainty equivalent). The remaining columns list the mean judged probabilities and normalized certainty equivalents for implicit disjunctions, explicit disjunctions and constituent events that were evaluated separately and summed. Asterisks and dagger indicate significance of differences between adjacent entries.

both studies were also asked to price risky prospects and complete a task that allowed us to estimate the shape of their value function for monetary gains.

Table 14.6 presents the mean judged probabilities and normalized certainty equivalents for participants in the studies of Fox and See (2002). The column labeled (A) presents the mean normalized certainty equivalent and judged probability for simple events, or "implicit disjunctions" in the language of support theory (for example, "interest rates go up"). The column labeled ($A_1 \vee A_2$) presents the mean normalized certainty equivalent and judged probability for the same events described as explicit disjunctions (for example, "interest rates go up and unemployment goes up, or interest rates go up and unemployment goes down"). The column labeled ($A_1 + A_2$) presents the mean sum of normalized certainty equivalents and judged probabilities for the same constituent events when assessed separately (for example, "interest rates go up and unemployment goes up"; "interest rates go up and unemployment goes down"). Judged probabilities reported in Table 14.6 accord perfectly with support theory, and certainty equivalents clearly violate both parts of the partition inequality (Equation (8)), consistent with the two-stage model.

Table 14.7 lists the proportion of participants in each of the Fox and See (2002) studies that violate various facets of the partition inequality (Equation (8)), as well as the proportion of participants whose judged probabilities follow an analogous pattern. The last column displays the proportion of participants whose certainty equivalents and judged probabilities exhibit subadditivity. Just as we saw in the studies summarized in Table 14.5 (see column 6), a large majority of participants in the Fox and See studies violated risk aversion (part (*ii*) of Equation (8)), exhibiting subadditivity of certainty equivalents. Moreover, a large majority of participants showed an analogous pattern of subadditivity among judged probabilities.

The fourth and fifth columns of Table 14.7 display tests of the more refined predictions of the partition inequality (Equation (8)). The results in column 4 show that a majority of participants reported higher certainty equivalents when events were described as explicit disjunctions (violating the first part of Equation (8)). A similar pattern appeared among judged probabilities (satisfying "implicit subadditivity" in the language of support theory; see Rottenstreich & Tversky, 1997). Finally, the results in column 5 show that a large majority of participants reported higher certainty equivalents when constituent prospects were evaluated separately and summed than when evaluating a single prospect whose target event was described as an explicit disjunction of the same constituent events (violating

Table 14.7 Percentage of participants violating the partition inequality

Study	N	Source of uncertainty		Implicit < Explicit $A < A_1 \vee A_2$	*Explicit < Sum* $A_1 \vee A_2 < A_1 + A_2$	*Implicit < Sum* $A < A_1 + A_2$
Fox & See (2002)						
Duke basketball fans	51	Basketball	$P(.)$	63%†	78%**	86%**
			$C(.)$	65%*	84%**	94%**
Brown students	29	Economic indicators	$P(.)$	69%*	66%†	83%**
			$C(.)$	59%	62%	66%†

† $p < .10$; * $p < .05$; ** $p < .005$.
The first three columns identify the participant population, sample size and sources of uncertainty. The fourth column indicates the relevant dependent variable (judged probability or normalized certainty equivalent). The remaining columns list the percentage of participants who violate the designated facet of the partition inequality (for certainty equivalents) and satisfy the designated prediction of support theory (for judged probability); asterisks and dagger indicate the statistical significance of the percentage.

the second part of Equation (8)). A similar pattern was exhibited for judged probabilities (satisfying "explicit subadditivity" in the language of support theory).

Fitting the Two-Stage Model

The two-stage model can be tested more directly against expected utility theory by comparing the fit of both models to empirical data. To fit the two-stage model, we did the following. For each target event, A, we observed its median judged probability $P(A)$. We next searched for the median certainty equivalent C of the risky prospect (x, p), where $p = P(A)$ (recall that these studies also asked participants to price risky prospects). Hence,

$C(x, A)$ is estimated by $C(x, p)$ where $p = P(A)$.

To illustrate, consider the study of basketball fans reported by Fox and See (2002). One of the prospects presented to these participants offered \$160 if Duke won its upcoming game against the University of North Carolina (UNC). To fit the two-stage model, we first observed that the median judged probability of the target event was .8. Next we found that the median certainty equivalent of this same population for the prospect that offered \$160 with probability .8 was \$119. The median certainty equivalent for the prospect that offered \$160 if Duke won its game against UNC was \$120. Thus, the error in this case was \$120−\$119 = \$1, or 0.6 percent.

To fit the classical theory, let C_A be the certainty equivalent of the prospect (\$160, A). Setting $u(0) = 0$, the classical theory yields $u(C_A) = u(160)\boldsymbol{P}(A)$, where u is concave and $\boldsymbol{P}(A)$ is an additive subjective probability measure.[19] Hence, $\boldsymbol{P}(A) = u(C_A)/u(160)$. Previous studies (e.g., Tversky, 1967; Tversky & Kahneman, 1992) have suggested that the utility function for small to moderate gains can be approximated by a power function: $v(x) = x^{\alpha}$, $\alpha > 0$. Thus, risk aversion (a concave utility function) is modeled by $\alpha < 1$, risk neutrality (a linear utility function) is modeled by $\alpha = 1$, and risk seeking (a convex utility function) is modeled by $\alpha > 1$. An independent test of risk attitudes of these same participants yielded an estimate of α, assuming expected utility theory.

[19] Note that the additive subjective probability measure $\boldsymbol{P}(.)$ should be distinguished from judged probability $P(.)$ and objective probability p.

Table 14.8 Comparison of data fit for two-stage model and classical theory

Study	N	α	Source of uncertainty	Mean absolute error		Superior fit of 2SM
				Two-stage model	Classical theory	
Fox & Tversky (1998)						
NBA fans	50	0.83	Basketball	0.04	0.15	90%**
Stanford students	82	0.80	Economic indicators	0.04	0.08	61%*
Fox & See (2002)						
Duke basketball fans	51	0.63	Basketball	0.13	0.18	74%**
Brown students	29	0.87	Economic indicators	0.12	0.43	97%**

* $p < .05$; ** $p < .01$.
The first two columns identify the participant population and sample size. The third column lists the median value of alpha (α) estimated for participants in each sample. The fourth column identifies the relevant source of uncertainty. The fifth and sixth columns indicate the mean absolute error in fitting median data to the two-stage model and classical theory, respectively. The final column lists the percentage of participants for whom the fit of the two-stage model was superior to the fit of the classical theory; asterisks indicate the statistical significance of the percentage.

As anticipated, participants in the studies of Fox and Tversky (1998) and Fox and See (2002) exhibited risk-averse utility functions under expected utility theory. Median estimates of α for each of these studies, listed in the fourth column of Table 14.8, range from 0.63 to 0.87. Moreover, a significant majority of participants in these studies exhibited $\alpha \leq 1$ (92 percent of NBA fans, 98 percent of Stanford University students, 94 percent of Duke University basketball fans and 83 percent of Brown University students; $p < .001$ for all samples). This confirms our earlier assertion that part (*ii*) of the partition inequality would be expected to hold for most participants under expected utility theory.

Subjective probabilities were estimated as follows. For each target event A, we computed $(C_A/160)^{\alpha}$ and divided these values by their sum to ensure additivity. Fits of both the two-stage model and classical theory for the studies of Fox and Tversky (1998) and also Fox and See (2002) are listed in Table 14.8. Based on the median certainty equivalents obtained in the studies listed, the fit of the two-stage model was consistently better than the fit of the classical theory. The mean absolute error for the two-stage model ranges from one-quarter to one-half of the mean absolute error of expected utility theory. Moreover, when this analysis was replicated for each participant, the two-stage model fit the data better than the classical theory for a significant majority of participants in all four studies.

14.2.5 From Allais to the Two-Stage Model: Summary

The Allais (1953) paradox and fourfold pattern of risk attitudes (see Table 14.4) suggest that consequences are not weighted by an additive probability measure. Instead, it seems that consequences are weighted by an inverse-S shaped transformation that overweights small probability consequences and underweights moderate to large probability consequences. This weighting function reflects diminished sensitivity to events between the natural boundaries of impossibility and certainty, which is formally expressed as *bounded subadditivity*. Numerous empirical studies have documented significant bounded subadditivity for chance bets (risk) and a more pronounced degree of subadditivity for bets contingent on natural events (uncertainty). This reduced sensitivity to uncertainty can be attributed to subadditivity of judged probability. According to the two-stage model (Tversky & Fox, 1995; Fox & Tversky, 1998), decision makers first judge the probability of the target event (consistent with support theory), and then weight that probability by an inverse-S shaped function (as in prospect theory). This two-stage model is a radical departure from previous models because it allows decision weights to violate both the assumptions of additivity and description invariance.

Evidence from a number of studies suggests that the two-stage model generally provides a better fit to the data than does the classical model. Moreover, both prospect theory and the two-stage model accommodate the Allais pattern in Problem 1, which reflects a preference for certainty. We now turn to the Ellsberg paradox illustrated in Problem 2. As we observed in Section 1, the Ellsberg pattern also violates the classical theory. However, the two-stage model cannot accommodate the Ellsberg paradox, because the two-stage model attaches the same weight to all events with probability p, regardless of their source. Hence, under the two-stage model, a prize that obtains with probability $1/3$ should be equally attractive, regardless of whether the probability is clear (for example, option E in Problem 2) or vague (option F). Similarly, a prize that obtains with probability $2/3$ should be equally attractive, regardless of whether the probability is clear (option H) or vague (option G). This prediction is contradicted by modal preferences exhibited in the Ellsberg problem.

In order to develop a more satisfactory account of decision under uncertainty, we must first gain a deeper understanding of the Ellsberg phenomenon: the empirical results that have been documented and how they might be interpreted. At that point, we will be ready to extend the two-stage model to accommodate this phenomenon.

14.3 THE PREFERENCE FOR KNOWLEDGE: FROM ELLSBERG TO THE COMPARATIVE IGNORANCE HYPOTHESIS

Recall that the violation of the sure-thing principle observed in the Allais paradox (Problem 1) was explained by diminished sensitivity to changes in probability away from zero and one. The violation observed in the Ellsberg paradox (Problem 2) resonates with a very different intuition: people prefer to bet on known rather than unknown probabilities. This interpretation is brought into sharper focus in a simpler, two-color problem that was also advanced by Ellsberg (1961). Imagine two urns, both containing red and black balls. Urn I contains 50 red balls and 50 black balls, whereas Urn II contains 100 red and black balls in an unknown proportion. Suppose that your task is to guess a color, and then draw a ball from one of the urns without looking. If you draw the color that you had guessed, you win a prize, say, \$10. Most people would rather bet on drawing a black ball from the known probability urn than a black ball from the unknown probability urn, and most people would likewise rather bet on drawing a red ball from the known probability urn than a red ball from the unknown probability urn. This pattern violates expected utility theory because it implies that the subjective probabilities for red and black are greater for the 50−50 urn than for the unknown probability urn and therefore cannot sum to one for both urns.

The Ellsberg problem has garnered much attention because in real-world contexts decision makers are seldom provided with precise probabilities of potential consequences. Ellsberg (1961, 2001) argued that willingness to act under uncertainty is governed not only by the perceived likelihood of target events and the attractiveness of potential consequences, but also the *ambiguity* of the information on which the likelihood judgment is based (that is, the degree of uncertainty concerning probabilistic information). Ellsberg observed that people generally find ambiguity aversive. More generally, we say that people prefer to bet on sources of uncertainty for which events have known rather than unknown probabilities.[20]

In this section, we begin with a review of the empirical study of ambiguity and source preference. Second, we describe ways in which researchers can establish that a decision maker prefers one source of uncertainty to another. Third, we discuss the psychological interpretation of ambiguity aversion and source preference. Finally, we outline ways in which source preference can be incorporated into the two-stage model that was developed in the previous section.

14.3.1 The Empirical Study of Ambiguity Aversion

Although Ellsberg presented no experimental evidence, the preference to bet on known rather than unknown probabilities has been demonstrated in numerous empirical studies using variations of Ellsberg's original problems (for a review of the literature on ambiguity aversion, see Camerer & Weber, 1992). In particular, a number of researchers have provided participants with either precise information concerning probabilities or no information concerning probabilities, as in the original Ellsberg problems. These studies provide empirical support for the predicted pattern of choices in the Ellsberg two-color problem (Raiffa, 1961; Becker & Brownson, 1964; Yates & Zukowski, 1976; Kahn & Sarin, 1988; Curley & Yates, 1989; Eisenberger & Weber, 1995) and three-color problem (Slovic & Tversky, 1974; MacCrimmon & Larsson, 1979).

Numerous studies have extended the exploration of ambiguity aversion by examining the effect of increasing the degree of second-order uncertainty (that is, uncertainty about probability). Traditionally, researchers have relied on four different methods for manipulating what they interpret to be "ambiguity". First, some experimenters have varied the width of a *range* of possible probabilities. For example, participants might be asked to price gambles with probabilities of (a) .5, (b) somewhere between .4 and .6, or (c) somewhere between .3 and .7 (Becker & Brownson, 1964; Curley & Yates, 1985, 1989; Kahn & Sarin, 1988). Second, some researchers have *endowed* participants with a probability and a qualitative degree of second-order uncertainty (Einhorn & Hogarth, 1986; Kahn & Sarin, 1988; Hogarth, 1989; Hogarth & Kunreuther, 1989; Hogarth & Einhorn, 1990; Kunreuther et al., 1995). For example, Kunreuther et al. (1995) manipulate ambiguity in the context of underwriter decision making by telling participants either that "all experts agree that the probability of a loss is [p]" or that the experts' best estimate of the probability of a loss is [p], but "there is wide disagreement about this estimate and a high degree of uncertainty among experts" (p. 11). Third, some researchers have provided participants with small and large random *samples* of information with identical proportions (Chipman, 1960; Beach & Wise, 1969; Gigliotti & Sopher, 1996). For instance, Chipman (1960) had participants choose between betting on a box with a known proportion of 100 match stems and heads versus

[20] Tversky and Wakker (1995, p. 1270) define sources of uncertainty as families of events that are assumed to be closed under union and complementation (that is, if events A_1 and A_2 are members of source **A**, then so are $A_1 \cup A_2$, $S-A_1$, and $S-A_2$). This assumption is satisfactory with one salient exception: in the Ellsberg three-color example (Problem 2), we might interpret known probability events {red, white ∪ blue} as one source of uncertainty and unknown probability events {white, blue, red ∪ white, red ∪ blue} as a second source of uncertainty. Hence, for our purposes, it is the fundamental character of the information on which judgment is based that defines a source.

a box with an unknown proportion of 100 stems and heads from which 10 items had been sampled (yielding a matched proportion of stems and heads). Finally, some researchers have manipulated ambiguity by providing participants with a *game of chance* entailing a multistage lottery in which the outcome probability is determined by a first-stage drawing (Yates & Zukowski, 1976; Larson, 1980; Bowen & Qui, 1992). For instance, in a study by Larson (1980), the proportion of winning poker chips in a container was to be determined by a number written on a card to be randomly drawn from a deck with 20 cards. Pairs of decks were constructed with fixed means and normal distributions surrounding those means; half the decks had a relatively large variance, and half had a relatively small variance.

Studies using the methods discussed above seem to provide broad empirical support for the notion that the attractiveness of a prospect generally decreases as second-order uncertainty increases. Many of these studies have also claimed to find "ambiguity seeking" for low probability gains. However, an important caveat is in order when evaluating the results of any study using one of these methods for manipulating ambiguity: they do not necessarily control for variations in subjective probability. Heath and Tversky (1991; see especially Table 4, p. 24) provide evidence that subjective probabilities associated with higher variance or less reliable estimates of p may be less extreme (that is, closer to .5) than those associated with lower variance or more reliable estimates of p. This pattern is also consistent with a model in which people anchor on an "ignorance prior" probability of one-half, and then adjust according to information provided them, with greater adjustment in response to more precise probabilistic information (cf. Fox & Rottenstreich, 2003). Hence, when participants learn that the best estimate of probability is .1 but "there is wide disagreement about [the] estimate and a high degree of uncertainty among experts", they may adopt a posterior probability that is higher than participants who learn that the best estimate is .1 and "all experts agree" (cf. Hogarth & Kunreuther, 1989). Such a pattern would mimic "ambiguity seeking" for low probability gains, but, in fact, it merely reflects a variation in subjective probability. In sum, although the studies using the foregoing methods (range, endowment, sampling and games of chance) make a persuasive case for ambiguity aversion in situations where there are clear and vague probabilities of .5, the interpretation of ambiguity seeking for low-probability gains should be regarded with some skepticism. In order to establish ambiguity aversion (or seeking), one must be careful to control for unintended variations in belief that may be introduced by the manipulation of ambiguity. Before turning to a discussion of how one might interpret these empirical observations, we must first address the question of how to distinguish ambiguity aversion—or more generally, the preference to bet on one source of uncertainty over another—from an account that can be attributed to differences in belief.

14.3.2 Establishing Source Preference

Two methods for establishing source preference can be defended, which we will call (1) probability matching and (2) complementary bets. Although we use games of chance (such as balls drawn from an urn) to illustrate these methods, the two methods allow us to extend the study of ambiguity to the domain of natural events, such as the future outcome of an election or future close of the stock market.

The probability matching method can be formalized as follows. Let **A** and **B** be two different sources of uncertainty. For instance, source **A** might be the color of a ball drawn from an urn containing 50 red and 50 black balls, whereas source **B** might be the color of a ball drawn from an urn containing 100 red and black balls in unknown proportion. A decision maker is said to prefer source **A** to source **B** if for any event A in **A** and B in **B**, $P(A) = P(B)$ implies $W(A) \geq W(B)$ or equivalently, $P(A) = P(B)$ implies $C(x, A) \geq C(x, B)$ for all $x > 0$. Thus, a person who says event A and event B are equally likely, but strictly prefers to bet on A, exhibits a preference for source **A** over source **B**. For instance, suppose a person says the probability of "black from the 50−50 urn" is .5 and the probability of "black from the unknown probability urn" is also .5. If this person prefers betting on black from the known probability urn to black from the unknown probability urn, then she has exhibited ambiguity aversion that cannot be readily attributed to differences in belief.

One could potentially object to the probability matching method on the grounds that it relies on an expression of judged probability rather than a measure of belief that is inferred from preferences. A second method for establishing source preference does not rely on judged probability. A decision maker is said to prefer source **A** to source **B** if for any event A in **A** and B in **B**, $W(A) = W(B)$ implies $W(S{-}A) \geq W(S{-}B)$, or equivalently, $C(x, A) = C(x, B)$ implies $C(x, S{-}A) \geq C(x, S{-}B)$ for all $x > 0$. Thus, a person who would rather bet on event A (for example, drawing a black ball from the 50−50 urn) than on event B (for example, drawing a black ball from the unknown probability urn) and would rather bet against A (that is, drawing a red ball from the 50–50 urn) than against B (that is, drawing a red ball from the unknown probability urn) exhibits source preference that cannot be readily attributed to differences in belief.[21]

14.3.3 Interpreting Ambiguity Aversion

Most of the aforementioned empirical studies of ambiguity aversion have manipulated ambiguity through vagueness of probabilities. Indeed, Ellsberg (1961) himself proposed to model ambiguity aversion in terms of probability vagueness (that is, the range of possible probabilities), and most models of ambiguity that followed Ellsberg have parameterized features of a second-order probability distribution (for a review, see Camerer & Weber, 1992, pp. 343–347). However, Ellsberg originally characterized ambiguity aversion not as vagueness aversion per se, but rather as reluctance to bet in situations where the decision maker perceives that he lacks adequate information or expertise:

> An individual . . . can always assign relative likelihoods to the states of nature. But how does he *act* in the presence of uncertainty? The answer to that may depend on another judgment, about the reliability, credibility, or adequacy of his information (including his relevant experience, advice and intuition) as a whole. (Ellsberg, 1961, p. 659)

In recent years, many behavioral researchers have returned to the original interpretation of ambiguity aversion as driven by the decision maker's confidence in his or her knowledge, skill or information (e.g., Frisch & Baron, 1988; Heath & Tversky, 1991). That is, ambiguity aversion might be attributed to reluctance to bet in situations where the decision maker feels relatively ignorant.

[21] This pattern could be attributed to differences in belief if the sum of probabilities of complementary events is lower for the less familiar source of uncertainty. Recall that support theory (Tversky & Koehler, 1994; Rottenstreich & Tversky, 1997) holds that judged probabilities of complementary events generally sum to one. Numerous studies provide evidence supporting this prediction (for a review of evidence, see Tversky & Koehler, 1994; Table 14.5 of this chapter). For counterexamples, see Brenner and Rottenstreich (1999); Macchi, Osherson and Krantz (1999) and Idson et al. (2001).

Vagueness Aversion Versus Ignorance Aversion

Under most circumstances, the ignorance versus vagueness conceptions are confounded: when a decision maker feels less knowledgeable, her judged probabilities are less precise. In order to tease apart these two accounts, we must find a circumstance in which vagueness aversion and ignorance aversion imply different patterns of choice. Heath and Tversky (1991) conducted a series of experiments comparing people's willingness to bet on their uncertain beliefs to their willingness to bet on chance events. Contrary to the vagueness-aversion hypothesis, Heath and Tversky found that people prefer to bet on their vague beliefs in situations where they feel especially knowledgeable or competent—though they prefer to bet on chance when they do not feel especially knowledgeable or competent. For instance, in one study, participants were asked to order their preferences among bets contingent on three sources of uncertainty: chance events, the winner of various professional football games and the winner of various states in the 1988 presidential election. Participants who rated their knowledge of football to be high and their knowledge of politics to be low preferred betting on football games to chance events that they considered equally probable. However, these participants preferred betting on chance events to political events that they considered equally probable. Analogously, participants who rated their knowledge of football to be low and their knowledge of politics to be high favored bets on politics to chance and chance to football.

The foregoing demonstration provides evidence for the ignorance-aversion hypothesis and casts doubt on the vagueness-aversion hypothesis. This demonstration relies on the probability matching method described in Section 14.3.2. In other studies using the complementary bets method, Heath and Tversky (1991) found that participants were willing to pay more to bet both for and against familiar events (for example, "more than 85 percent of undergraduates at [your university] receive on-campus housing") than for or against matched events that were less familiar (for example, "more than 70 percent of undergraduates at [a less familiar university] receive on-campus housing"). The preference to bet on more familiar events has since been replicated by a number of researchers using both the probability matching method (Taylor, 1995; Taylor & Kahn, 1997) and the complementary bets method (Keppe & Weber, 1995; Tversky & Fox, 1995).

The perspective we are advancing is that the ambiguity-aversion phenomenon is driven by the decision maker's perception of her level of knowledge concerning the target event, rather than by features of the second-order probability distribution. Heath and Tversky (1991) provided support for this interpretation by identifying situations where decision makers preferred to bet on their vague assessments of familiar events rather than bet on chance events with matched probability. We are not arguing that perceptions of one's own competence influence probability vagueness or that vagueness influences perceived competence. Rather, characteristics of a source of uncertainty influence both the precision with which likelihood can be assessed by the decision maker and the decision maker's subjective perception of her own competence judging likelihood. However, it is the perception of competence that drives willingness to act under uncertainty. Indeed, Heath and Tversky (1991) found that two-thirds of participants preferred to bet on their guess of whether a randomly picked stock would go up or down the next day rather than bet on their guess of whether a randomly picked stock had gone up or down the previous day. Clearly, the vagueness of one's judgment is unlikely to be influenced by whether the stock is picked from yesterday's or tomorrow's paper (if anything, participants could have more precise knowledge of yesterday's stock movement). However, when betting on yesterday's close, participants may feel relatively ignorant because they can already be wrong at the time of "postdiction" (see also Brun & Teigen, 1990).

The Comparative Ignorance Hypothesis

The ignorance-aversion hypothesis asserts that source preference is driven by the decision maker's subjective appraisal of his or her knowledge concerning target events rather than some second-order measure of probability vagueness. Fox and Tversky (1995) extended this account by asking what conditions produce ignorance aversion. They conjectured that a decision maker's confidence betting on a target event is enhanced (diminished) when he contrasts his knowledge of the event with his inferior (superior) knowledge about another event, or when he compares himself with less (more) knowledgeable individuals. According to the "comparative ignorance hypothesis", ambiguity aversion is driven by a comparison with more familiar sources of uncertainty or more knowledgeable people and is diminished in the absence of such a comparison. Three nuances of this account are worth emphasizing: (1) source preference increases with the salience of contrasting states of knowledge; (2) source preference is relative rather than absolute; (3) source preference is a function of decision makers' appraisal of their relative *knowledge* rather than their *information*.

The Salience of Contrasting States of Knowledge

Virtually every empirical study of ambiguity aversion reported before Fox and Tversky (1995) relied on a within-subject design in which all participants evaluated multiple sources of uncertainty (for example, in which each participant priced bets drawn from both clear and vague probability urns). In a series of experiments, Fox and Tversky (1995) documented pronounced ambiguity aversion in *comparative* contexts in which each participant evaluated lotteries with both clear and vague probabilities, but they found the effect was greatly diminished—or disappeared entirely—in *noncomparative* contexts in which different groups of participants evaluated the lotteries in isolation. For instance, participants in a comparative condition said they were willing to pay $24.34, on average, for a bet that offered $100 if they correctly guessed the color drawn from an urn containing 50 red balls and 50 black balls, but they would pay only $14.85, on average, for a bet that offered $100 if they correctly guessed the color drawn from an urn containing 100 red and black balls in an unknown proportion. In contrast, participants who priced the 50–50 bet in isolation were willing to pay $17.95, on average, whereas participants who priced the unknown probability bet in isolation were willing to pay $18.42, on average. Hence, the Ellsberg result seemed to disappear when the experiment was run as a between-subject design rather than a within-subject design. A similar pattern was observed in a follow-up study in which participants evaluated prospects contingent on the future temperature in San Francisco (a familiar city) and/or Istanbul (an unfamiliar city). Chow and Sarin (2001) replicated the finding that source preference greatly diminishes in noncomparative contexts, though in some of their studies it did not disappear entirely. Further evidence from market studies has shown that the pronounced difference in prices for clear versus vague bets, observed when both bets are traded together, diminishes or disappears when these bets are traded in separate markets (Sarin & Weber, 1993).

The comparative ignorance hypothesis asserts that source preference is driven by the salience of contrasting states of knowledge. It applies not only to the contrast in a decision maker's knowledge of two different sources of uncertainty but also to the contrast between a decision maker's knowledge and that of other people. Fox and Tversky (1995) showed that participants who were told that more knowledgeable people would be making the same choice were much less likely to bet on their prediction than participants who were not told about such experts. For instance, in one study, psychology undergraduates were asked whether they thought that a particular stock would close higher in its next day of trading, and then asked whether they would prefer to receive $50 for sure or $150 if their prediction was correct. In this case, most favored the uncertain prospect. However, when a second group of participants was told that the survey was also being presented to economics graduate students and professional stock analysts, most favored the sure payment. In a similar vein, Chow and Sarin (1999) found that the preference for known over unknown probabilities is amplified when participants pricing a single bet (for example, a draw from a 50–50 urn) are made aware of the fact that another participant or group of participants is evaluating a different bet that offers a contrasting degree of information about probabilities (for example, a draw from an unknown probability urn).

Relative Versus Absolute Knowledge

Fox and Tversky (1995) emphasized that the distinction between comparative and non-comparative assessment refers to the state of mind of the decision maker rather than the experimental context. The studies cited above facilitated such comparisons by juxtaposing a known probability or familiar prospect with an unknown probability or unfamiliar prospect, or by explicitly mentioning other people who were more (or less) knowledgeable. More recently, Fox and Weber (2002) have shown that the comparative state of mind that drives source preference can be manipulated without resorting to this "joint–separate" evaluation paradigm. For instance, participants in one study were less willing to bet on their prediction of the outcome of the upcoming Russian election (a modestly familiar event) when they had been previously asked a question concerning the upcoming US election (a more familiar event) than when they had been previously asked a question concerning the upcoming Dominican Republic election (a less familiar event). This demonstration also provides evidence that when a state of comparative ignorance is induced (by reminding people of alternative states of knowledge), willingness to act is governed by the decision maker's *relative* knowledge judging the target event rather than some measure of *absolute* knowledge.[22]

Knowledge Versus Information

Fox and Weber (2002) also demonstrated that comparisons can be facilitated even in a decision context that is not explicitly comparative. In particular, when people are provided with information they do not know how to use, this may remind them of their ignorance relative to experts who do know how to use that information. For instance, in one study, psychology students were less willing to bet on their assessment of the inflation rate in the Netherlands if they had been provided information concerning the country's gross domestic product

[22] For a demonstration in the context of strategic uncertainty, see also Fox and Weber (2002), study 5.

growth, unemployment rate and prevailing interest rate than if they had been provided no such information. This demonstration also suggests that comparative ignorance effects are governed by the decision maker's perception of her relative *knowledge* concerning the target event rather than the absolute amount of relevant *information* that she has available.

14.3.4 Incorporating Source Preference into the Two-Stage Model

The comparative ignorance effect presents serious modeling challenges. In particular, demonstrations using the joint–separate evaluation paradigm show that strict source preference sometimes disappears when prospects are evaluated separately, rather than jointly. This violates the principle of *procedure invariance*, according to which strategically equivalent elicitation procedures should produce the same preference ordering (Tversky, Sattath & Slovic, 1988; see also Tversky, 1996). More troubling still is the problem that the comparative ignorance phenomenon is inherently subjective and context-dependent. To predict source preference *ex ante*, one must somehow parameterize (1) the decision maker's sense of his or her relative knowledge regarding the events in question and (2) the salience of alternative states of knowledge.

These caveats notwithstanding, it would certainly be valuable to incorporate source preference into the two-stage model. Recall that according to the two-stage model, a decision maker first judges probabilities, consistent with support theory, and then transforms these probabilities in a way that accords with his or her weighting of chance events. Hence, the original specification of the two-stage model does not allow for source preference. Although the model presented in Equation (5) may provide a reasonable first-order approximation of people's decision behavior under uncertainty, it is important to note that this specification will fail when source preference is especially pronounced.

Fox and Tversky (1998) acknowledged this limitation and proposed that their model could be extended to accommodate source preference while preserving the segregation of belief (that is, judged probability) and preference (that is, decision weights). They suggested generalizing Equation (5) by letting $W(A) = F[P(A)]$, so that the transformation F of probability depends on the source of uncertainty. They assume that F, like the risky weighting function, w, is a subadditive transformation of P (see Tversky & Fox, 1995; Tversky & Wakker, 1995; Wakker, in press). One convenient parameterization may be defined by

$$W(A) = (w[P(A)])^{\theta}, \tag{9}$$

where $\theta > 0$ is inversely related to the attractiveness of the particular source of uncertainty. A second approach is to vary a parameter of the probability weighting function that increases weights throughout the unit interval. For instance, one could vary the parameter δ of Prelec's (1998) two-parameter weighting function, $w(p) = \exp(-\delta\ (-\ln p)^{\gamma})$, where $\delta > 0$ is inversely related to the attractiveness of the source. This latter scheme has the advantage of manipulating "elevation" (that is, source preference) independently of the degree of "curvature" (that is, source sensitivity). It also allows for the possibility of differences in sensitivity to probabilities drawn from different sources of uncertainty that can be modeled through changes in the parameter γ. Such a decrement in sensitivity was recently documented for an extremely unfamiliar domain by Kilka and Weber (2001).

An approach that allows for differences in curvature of probability weights may have a second application. The two-stage model, like most normative and descriptive models of decision under uncertainty, presumes that (weighted) beliefs and (the subjective value of) consequences can be segregated into separate terms that are independent of one another. Recently, the generality of even this basic principle has been called into question. Rottenstreich and Hsee (2001) observed that the weighting function may exhibit greater curvature for more "affect-rich" outcomes, such as the possibility of receiving an electrical shock or a kiss from a favorite movie star, than for "affect-poor" outcomes, such as money. For instance, these researchers report that their participants found an electric shock about as unattractive as a penalty of $20. However, participants were willing to pay seven times as much to avoid a 1 percent chance of an electric shock as they were to avoid a 1 percent chance of losing $20. This phenomenon begs further study—in particular, a more tightly circumscribed definition and independent measure of "affect richness" would be useful. In any case, the phenomenon can easily be accommodated by the extended two-stage model if we allow the curvature parameter of the probability weighting function to vary with the degree of affect richness of the target outcome.

14.3.5 From Ellsberg to the Comparative Ignorance Hypothesis: Summary

The Ellsberg paradox and similar demonstrations have established that willingness to act under uncertainty depends not only on the degree of uncertainty but also on its source. The empirical study of decision under uncertainty has suggested that the aversion to ambiguity does not reflect a reluctance to bet on vaguer probabilities, but rather a reluctance to act when the decision maker feels less knowledgeable. Moreover, it appears that the awareness of relative ignorance occurs only to the extent that contrasting states of knowledge are especially salient to the decision maker, an account known as the "comparative ignorance hypothesis". Source preference can be accommodated by the two-stage model through a generalization that weights judged probabilities by a subadditive function F, the elevation of which depends on the source of uncertainty.

14.4 SUMMING UP: A BEHAVIORAL PERSPECTIVE ON DECISION UNDER UNCERTAINTY

The Allais and Ellsberg paradoxes present challenges to the classical theory of decision under uncertainty that are so robust and so fundamental that they appear in most introductory texts of microeconomics (e.g., Kreps, 1990), decision analysis (e.g., von Winterfeldt & Edwards, 1986) and behavioral decision theory (e.g., Baron, 2001). The implications of these anomalies continue to preoccupy decision theorists to this day. Although modal preferences in both problems violate the sure-thing principle, the violations seem to reflect a distinct psychological rationale: the preference for certainty (Allais) versus the preference for knowledge (Ellsberg). Nevertheless, both patterns can be accommodated by a nonlinear weighting function that models source sensitivity through its degree of curvature and source preference through its elevation.

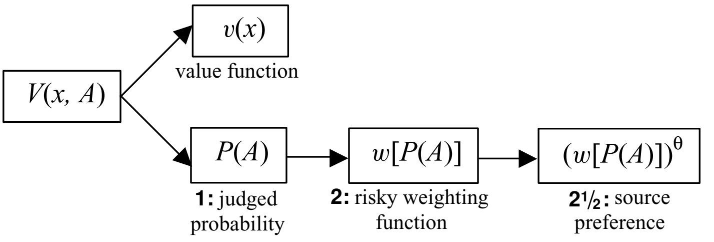

Figure 14.7 Visual depiction of the extended two-stage model, as parameterized in Equation (9). $V(x, A) = v(x)W(A) = v(x)(w[P(A)])^{\theta}$, where $V(x, A)$ is the value of the prospect that pays \$x if event A obtains (and nothing otherwise), $v(.)$ is the value function for monetary gains, $P(.)$ is judged probability, $w(.)$ is the risky weighting function, and θ is the source preference parameter

The two-stage model bridges two of the most influential strains of research in behavioral decision theory. It integrates the analysis of decision under risk and uncertainty articulated in prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) with work on heuristics and biases in judgment under uncertainty (e.g., Kahneman, Slovic & Tversky, 1982) that can be formalized through support theory (Tversky & Koehler, 1994; Rottenstreich & Tversky, 1997). The resulting model extends prospect theory by teasing apart the role of belief (that is, judged probability) and preference (that is, source sensitivity and source preference) in the weighting of events. It departs markedly from previous models by allowing decisions to differ depending on the way in which target events are described.

14.4.1 Value, Belief and Preference in Decision Under Uncertainty

We began this chapter with the observation that values and beliefs are key inputs in choice under uncertainty. However, we have identified systematic deficiencies of the classical approach in which values and beliefs are neatly segregated into utilities and subjective probabilities. In particular, we have observed that the weight decision makers afford uncertain events can be decomposed into judged probabilities that are weighted according to an inverse-S shaped function, the curvature of which reflects the decision maker's sensitivity to the particular source of uncertainty, and the elevation of which reflects the decision maker's preference to act on that source. A summary of our perspective on the role of value, belief and preference in decision under uncertainty, according to the extended two-stage model (as parameterized in Equation (9)), is presented in Table 14.9 and illustrated in Figure 14.7.

First, willingness to act under uncertainty is influenced in this model by the subjective value of the target outcome, $v(x)$. This value function is characterized by *reference dependence:* that is, it is a function of losses and gains relative to a reference point (usually, the status quo) rather than absolute states of wealth as in expected utility theory. The function is *concave for gains*, giving rise to some risk aversion for gains, just as in expected utility theory (see Figure 14.1). However, the function is also *convex for losses*, giving rise to some risk seeking for losses, contrary to expected utility theory (see Figure 14.2a). Finally, the function is *steeper for losses than for gains*, giving rise to "loss aversion" that can appear

Table 14.9 Summary of the role of value, belief and preference in decision under uncertainty according to the extended two-stage model[a]

Function/ Parameter	Interpretation	Characteristics	Empirical implications	Key references
v(.)	Value function	1) Reference dependence	1) Framing effects	Kahneman & Tversky (1984)
		2) Concave for gains; convex for losses	2) Risk seeking for gains; risk aversion for losses	Tversky & Kahneman (1986) Kahneman, Knetsch & Thaler (1990)
		3) Steeper for losses than gains	3) Loss aversion	Tversky & Kahneman (1991)
P(.)	Judged probability	1) Subadditivity	1) $P(A) \leq P(A_1) + P(A_2)$	Tversky & Koehler (1994)
		2) Description-dependence	2) $P(A) \leq P(A_1 \vee A_2)$	Rottenstreich & Tversky (1997)
w(.)	Risky weighting function	1) Bounded subadditivity	1) Fourfold pattern of risk attitudes	Tversky & Fox (1995)
		2) Subcertainty	2) More risk aversion than risk seeking	Tversky & Wakker (1995) Wu & Gonzalez (1996) Prelec (1998) Gonzalez & Wu (1999)
θ	Source preference parameter	Source preference varies with salience and nature of contrasting states of knowledge	Comparative Ignorance Effect	Heath & Tversky (1991) Fox & Tversky (1995) Fox & Weber (2002)

[a] Fox & Tversky (1998); see also Tversky & Kahneman (1992); Kahneman & Tversky (1979).

as risk aversion.[23] For instance, most people would reject a bet that offered a .5 chance of winning $100 and a .5 chance of losing $100, because a potential loss of $100 has a greater psychological impact than a potential gain of the same amount. In fact, most people would require a .5 chance of gaining at least $200 to compensate for a .5 chance of losing $100.

Second, willingness to act under uncertainty is influenced in this model by the perceived likelihood that the target event will occur, which can be quantified as a judged probability, $P(A)$. Judged probabilities are characterized by *subadditivity*: the probability of an uncertain event is generally less than the sum of probabilities of constituent events. This may contribute to violations of the partition inequality that mimic risk-seeking behavior. Moreover, judged probabilities are *description dependent:* as the description of the target event is unpacked into an explicit disjunction of constituent events, judged probability may increase. This phenomenon gives rise, for example, to a greater willingness to pay for an

[23] In fact, risk aversion over modest stakes for mixed (gain-loss) prospects cannot be explained by a concave utility function because it implies an implausible degree of risk aversion over large stakes. See Rabin (2000).

insurance policy that covers hospitalization due to "accident or disease" than a policy that covers hospitalization for "any reason" (Johnson et al., 1993).

Third, willingness to act under uncertainty is influenced in this model by the risky weighting function, $w(.)$. The curvature of this weighting function provides an index of a person's diminishing sensitivity to changes in probability between the natural boundaries of zero and one, and the elevation of this weighting function provides an index of a person's overall risk preference. The risky weighting function is generally characterized by *bounded subadditivity:* a probability p has more impact on decisions when added to zero or subtracted from one than when added to or subtracted from an intermediate probability q.

This inverse-S shaped weighting function implies a fourfold pattern of risk attitudes: people tend to be risk averse for gains and risk seeking for losses of moderate to high probability (underweighting of moderate to high probabilities reinforces the risk attitudes implied by the shape of the value function), but they tend to be risk seeking for gains and risk averse for losses of low probability (overweighting of low probabilities reverses the risk attitudes implied by the shape of the value function).

The risky weighting function is also generally characterized by *subcertainty:* $w(p) + w(1-p) \leq 1$. Visually, subcertainty manifests itself as a weighting function that crosses the identity line below .5, and weighting functions with lower elevation are associated with more pronounced degrees of subcertainty. It is important to note that there are large individual differences in both the curvature and elevation of measured weighting functions, so that some people might exhibit pronounced bounded subadditivity and no subcertainty, some might exhibit little subadditivity and pronounced subcertainty, and others might exhibit both or neither (for a detailed investigation, see Gonzalez & Wu, 1999).

Finally, willingness to act under uncertainty is influenced in this model by the source preference parameter, θ. Source preference reflects a person's eagerness or reluctance to act on one source of uncertainty (for example, invest in the domestic stock market) rather than another (for example, invest in a foreign stock market). Decision makers prefer to act in situations where they feel relatively knowledgeable or competent to situations in which they feel relatively ignorant or incompetent, but only to the extent that comparative states of knowledge are salient. We suspect that the original formulation of the two-stage model is adequate in most environments, where a contrasting state of knowledge does not occur to the decision maker. Indeed, the research reviewed in this chapter suggests that the perception of comparative ignorance, when it does occur, is subjective, transitory and context-dependent.

14.4.2 Judged Probability Versus Subjective Probability

In the classical model of decision under uncertainty, direct judgments of probability are rejected in favor of a measure of belief that is inferred from choices between bets. This approach yields an elegant axiomatic theory that allows for the simultaneous measurement of utility and subjective probability. Unfortunately, the price of this parsimony is a set of restrictions on subjective probability that are descriptively invalid. In particular, the data reviewed in this chapter suggest that outcomes are weighted in a manner that violates both the assumptions of additivity and description invariance. Perhaps it was the observation of incoherent judged probabilities that led early theorists to reject direct expressions of belief in their development of a normative theory of choice. However, empirical studies reviewed in this chapter demonstrate persuasively that judged probabilities are, in fact, diagnostic of choices and therefore can be fruitfully incorporated into a *descriptive* model of decision under uncertainty. Moreover, support theory provides a formal framework for predicting departures from additivity and description invariance in judged probability. Thus, the descriptive theory of decision making under uncertainty may be informed and facilitated by a deeper understanding of judgment under uncertainty.[24]

14.4.3 Concluding Comments

In this chapter, we have reviewed a descriptive account of decision making under uncertainty. The picture that has emerged accommodates a wide range of empirical data that have been collected in recent decades and explains several puzzles in the literature. The (extended) two-stage model teases apart the role of value, belief and preference underlying willingness to act under uncertainty. Of course, the story presented here is incomplete, and a great deal of continuing research is called for. The present account accommodates the observation that real-world decisions require people to judge probabilities for themselves, with some degree of second-order uncertainty. However, we have confined most of the present discussion to prospects involving a single positive outcome, such as a bet that pays a fixed prize if the home team wins a sporting event (and nothing otherwise). Decisions in the real world often involve a host of multiple potential outcomes, some of which may entail losses. Although the cumulative version of prospect theory accommodates multiple uncertain outcomes entailing both gains and losses, further research is needed to determine an appropriate decomposition of decision weights into belief and preference terms in rank- and sign-dependent models (for an early attempt, see Wu & Gonzalez, 1999b).

Another promising avenue for future research is to develop tools for helping decision makers make more rational decisions. The present account suggests that several violations of rational choice theory can be avoided if the decision maker merely calculates the expected value of each option and is willing to bind her actions to the principle of expected value maximization. This simple procedure will eliminate inconsistencies in: (1) risk preference attributed to nonlinearity of the value function v; (2) risk preference attributed to nonlinearity of the risky weighting function w; (3) source preference attributed to the comparative ignorance effect and modeled by the parameter θ. However, the present account suggests that a final source of departures from rational choice—those attributed to incoherence of belief—are not so readily purged. The aforementioned study of options traders (Fox, Rogers & Tversky, 1996) is especially relevant in this respect. Options traders in this sample did indeed price chance prospects by their expected value, based on objective probabilities provided by the experimenter. Likewise, these experts priced uncertain prospects by their *apparent* expected value, based on probabilities they had judged for themselves. However, because these judged probabilities exhibited subadditivity, selling prices for uncertain prospects were also subadditive. It appears that even experts whose careers depend on their ability to make rational decisions under uncertainty have difficulty avoiding subadditivity in

[24] Note that according to the two-stage model, applying the inverse risky weighting function to uncertain decision weights will recover judged probabilities; that is, $w^{-1}(W(A)) = P(A)$, where w^{-1} is the inverse of w (cf. Wakker, in press).

their judgment and decision making. Because belief-based departures from rational choice cannot be eliminated by a simple act of volition, future prescriptive work might seek a better understanding of the psychological mechanisms underlying belief formation so that appropriate corrective measures might be developed.

REFERENCES

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science*, *11*, 1497–1512.
Allais, M. (1953). Le comportement de l'homme rationnel devant le risque, critique des postulates et axiomes de l'école américaine. *Econometrica, 21*, 503–546.
Anscombe, F.J. & Aumann, R.J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, *34*, 199–205.
Ayton, P. (1997). How to be incoherent and seductive: Bookmakers' odds and support theory. *Organizational Behavior and Human Decision Processes*, *72*, 99–115.
Baron, J. (2001). *Thinking and Deciding* (3rd edn). Cambridge: Cambridge University Press.
Beach, L.R. & Wise, J.A. (1969). Subjective probability and decision strategy. *Journal of Experimental Psychology*, *79*, 133–138.
Becker, S.W. & Brownson, F.O. (1964). What price ambiguity? On the role of ambiguity in decision making. *Journal of Political Economy*, *72*, 2–73.
Bleichrodt, H.B. & Pinto, J.L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science*, *46*, 1485–1496.
Bleichrodt, H., Pinto, J.L. & Wakker, P.P. (2001). Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Management Science*, *46*, 1498–1514.
Borel, É. (1924/1964). Apropos of a treatise on probability. In H. Kyburg & H. Smokler (eds), *Studies in Subjective Probability* (pp. 46–60). New York: Wiley.
Bowen, J. & Qui, Z. (1992). Satisficing when buying information. *Organizational Behavior and Human Decision Processes*, *51*, 471–481.
Brenner, L.A. & Y. Rottenstreich (1999). Focus, repacking, and judgment of grouped hypotheses. *Journal of Behavioral Decision Making*, *12*, 141–148.
Brun, W. & Teigen, K. (1990). Prediction and postdiction preferences in guessing. *Journal of Behavioral Decision Making*, *3*, 17–28.
Camerer, C.F. & Ho, T.H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, *8*, 167–196.
Camerer, C.F. & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*, 325–370.
Chipman, J.S. (1960). Stochastic choice and subjective probability. In D. Willner (ed.), *Decision, Values, and Groups* (vol. 1, pp. 70–85). Oxford, UK: Pergamon.
Chow, C.C. & Sarin, R.K. (2001). Comparative ignorance and the Ellsberg paradox. *Journal of Risk and Uncertainty*, *2*, 129–139.
Chow, C.C. & Sarin, R.K. (1999). Choices under uncertainty: Known, unknown, and unknowable. Unpublished paper, University of California at Los Angeles.
Cohen, M., Jaffray, J.Y. & Said, T. (1985). Individual behavior under risk and under uncertainty: An experimental study. *Theory and Decision*, *18*, 203–228.
Curley, S.P. & Yates, J.F. (1985). The center and range of the probability interval as factors affecting ambiguity preferences. *Organizational Behavior and Human Decision Processes*, *36*, 272–287.
Curley, S.P. & Yates, J.F. (1989). An empirical evaluation of descriptive models of ambiguity reactions in choice situations. *Journal of Mathematical Psychology*, *33*, 397–427.
deFinetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. English translation in H.E. Kyburg & H.E. Smokler (eds) [1964], *Studies in Subjective Probability* (pp. 93–158). New York: Wiley.
DeGroot, M.H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.

Edwards, W.D. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors*, *4*, 59–73.

Einhorn, H.J. & Hogarth, R.M. (1986). Decision making under ambiguity. *Journal of Business*, *59*, 225–249.

Eisenberger, R. & Weber, M. (1995). Willingness-to-pay and willingness-to-accept for risky and ambiguous lotteries. *Journal of Risk and Uncertainty*, *10*, 223–233.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.

Ellsberg, D. (2001). *Risk, Ambiguity and Decision*. New York: Garland.

Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 330–344.

Fishburn, P. & Kochenberger, G. (1979). Two-piece Von Neumann–Morgenstern utility functions. *Decision Sciences*, *10*, 503–518.

Fox, C.R. (1999). Strength of evidence, judged probability, and choice under uncertainty. *Cognitive Psychology*, *38*, 167–189.

Fox, C.R. & Birke, R. (2002). Lawyers exhibit subadditivity when forecasting trial outcomes. *Law and Human Behavior*, *26*, 159–173.

Fox, C.R. & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science*, *13*, 195–200.

Fox, C.R. & See, K.E. (2002). Testing the two-stage model in decision under uncertainty. Unpublished data, Fuqua School of Business, Duke University.

Fox, C.R. & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics*, *110*, 585–603.

Fox, C.R. & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science*, *44*, 879–895.

Fox, C.R. & Wakker, P.P. (2000). Value elicitation in Fox, Rogers & Tversky (1996): The method is flawed but the results are robust. Unpublished paper, University of Amsterdam.

Fox, C.R., Rogers, B.A. & Tversky, A. (1996). Options traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty*, *13*, 5–17.

Fox, C.R. & Weber, M. (2002). Ambiguity aversion, comparative ignorance, and decision context. *Organizational Behavior and Human Decision Processes*, *88*, 476–498.

Friedman, M. & Savage, L.J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, *56*, 279–304.

Frisch D. & Baron, J. (1988). Ambiguity and rationality. *Journal of Behavioral Decision Making*, *1*, 149–157.

Gigliotti, G. & Sopher, B. (1996). The testing principle: Inductive reasoning and the Ellsberg paradox. *Thinking and Reasoning*, *2*, 33–49.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, *16*, 65–88.

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

Heath, C. & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, *4*, 5–28.

Hershey, J.C. & Schoemaker, P.J. (1980). Prospect theory's reflection hypothesis: A critical examination. *Organizational Behavior and Human Decision Processes*, *25*, 395–418.

Hogarth, R.M. (1989). Ambiguity and competitive decision making: Some implications and tests. *Annals of Operations Research*, *19*, 31–50.

Hogarth, R.M. & Einhorn, H.J. (1990). Venture theory: A model of decision weights. *Management Science*, *36*, 780–803.

Hogarth, R.M. & Kunreuther, H. (1989). Risk, ambiguity, and insurance. *Journal of Risk and Uncertainty*, *2*, 5–35.

Idson, L.C., Krantz, D.H., Osherson, D. & Bonini, N. (2001). The relation between probability and evidence judgment: An extension of support theory. *Journal of Risk and Uncertainty*, *22*, 227–249.

Johnson, E.J., Hershey, J., Meszaros, J. & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, *7*, 35–51.

Kahn, B.E. & Sarin, R.K. (1988). Modeling ambiguity in decisions under uncertainty. *Journal of Consumer Research*, *15*, 265–272.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *4*, 263–291.

Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341–350.

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

Kahneman, D., Slovic, P. & Tversky, A. (eds) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Karni, E. & Mongin, P. (2000). On the determination of subjective probability by choices. *Management Science*, *46*, 233–248.

Keppe, H.J. & Weber, M. (1995). Judged knowledge and ambiguity aversion. *Theory and Decision*, *39*, 51–77.

Kilka, M. & Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty? *Management Science*, *47*, 1712–1726.

Knight, F.H. (1921). *Risk, Uncertainty, and Profit*. Boston, MA: Houghton-Mifflin.

Koehler, D.J. (1996). A strength model of probability judgments for tournaments. *Organizational Behavior and Human Decision Processes*, *66*, 16–21.

Koehler, D.J. (2000). Probability judgment in three-category classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 28–52.

Kreps, D.M. (1990). *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.

Kunreuther, H., Meszaros, J., Hogarth, R.M. & Spranca, M. (1995). Ambiguity and underwriter decision processes. *Journal of Economic Behavior and Organization*, *26*, 337–352.

Larson, J.R., Jr. (1980). Exploring the external validity of a subjectively weighted utility model of decision making. *Organizational Behavior and Human Performance*, *26*, 293–304.

Luce, R.D. (2000). *Utility for Gains and Losses: Measurement-Theoretical and Experimental Approaches.* New York: Erlbaum.

Luce, R.D. & Fishburn, P.C. (1992). A note on deriving rank-dependent utility using additive finite first-order gambles. *Journal of Risk and Uncertainty*, *11*, 5–16.

Macchi, L., Osherson, D. & Krantz, D. (1999). A note on superadditive probability judgment. *Psychological Review*, *106*, 210–214.

MacCrimmon, K.R. (1968). Descriptive and normative implications of the decision-theory postulates. In K. Borsch & J. Mossin (eds), *Risk and Uncertainty* (pp. 3–32). London: Macmillan.

MacCrimmon, K.R. & Larsson, S. (1979). Utility theory: Axioms versus "paradoxes". In M. Allais & O. Hagen (eds), *Expected Utility and the Allais Paradox* (pp. 27–145). Dordrecht: D. Reidel.

Machina, M.J. & Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica*, *60*, 745–780.

March, J.G. & Shapira, Z. (1987). Managerial perspectives on risk and risk-taking. *Management Science*, *33*, 1404–1418.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, *60*, 151–158.

Payne, J.W., Laughhunn, D.J. & Crum, R. (1981). Further tests of aspiration level effects in risky choice. *Management Science*, *27*, 953–958.

Pratt, J.W., Raiffa, H. & Schlaifer, R.O. (1964). The foundations of decision under uncertainty: An elementary exposition. *Journal of the American Statistical Association*, *59*, 353–375.

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

Preston, M. & Baratta, P. (1948). An experimental study of the auction-value of an uncertain outcome. *American Journal of Psychology*, *61*, 183–193.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, *3*, 323–343.

Rabin, M. (2000). Risk aversion and expected utility theory: A calibration theorem. *Econometrica*, *68*, 1281–1292.

Raiffa, H. (1961). Risk, ambiguity, and the Savage axioms: Comment. *Quarterly Journal of Economics*, *75*, 690–694.

Ramsey, F.P. (1931). Truth and probability. In R.B. Braithwaite (ed.), *The Foundations of Mathematics and Other Logical Essays by F.P. Ramsey* (pp. 158–198). New York: Harcourt, Brace.
Redelmeier, D.A., Koehler, D.J., Liberman, V. & Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified possibilities. *Medical Decision Making*, *15*, 227–230.
Rottenstreich, Y. & Hsee, C.K. (2001). Money, kisses, and electric shocks: An affective interpretation of probability weighting function. *Psychological Science*, *12*, 185–190.
Rottenstreich, Y. & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, *2*, 406–415.
Sarin, R.K. & Weber, M. (1993). Effects of ambiguity in market experiments. *Management Science*, *39*, 602–615.
Savage, L.J. (1954). *The Foundation of Statistics*. New York: Wiley.
Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*, 571–587.
Shafer, G. (1986). Savage revisited. *Statistical Science*, *1*, 463–501.
Sloman, S.A., Rottenstreich, Y., Hadjichristidis, C., Steibel, J.M. & Fox, C.R. (2003). Unpacking implicit probability judgment. Working paper, Brown University, Department of Cognitive and Linguistic Sciences.
Slovic, P. & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Science*, *19*, 368–373.
Taylor, K. (1995). Testing credit and blame attributions as explanation for choices under ambiguity. *Organizational Behavior and Human Decision Processes*, *54*, 128–137.
Taylor, K. & Kahn, B. (1997). An investigation of consumer reactions to ambiguity: The effects of knowledge, control, and accountability. University of Pennsylvania, Wharton School, Working Paper 96–006.
Teigen, K.H. (1974). Subjective sampling distributions and the additivity of estimates. *Scandinavian Journal of Psychology*, *24*, 97–105.
Tversky, A. (1967). Utility theory and additivity analysis of risky choices. *Journal of Experimental Psychology*, *75*, 27–36.
Tversky, A. (1996). Contrasting rational and psychological principles of choice. In R.J. Zeckhauser, R.L. Keeney & J.K. Sebenius (eds), *Wise Choices* (pp. 5–21). Boston, MA: Harvard Business School Press.
Tversky, A. & Fox, C.R. (1995). Weighing risk and uncertainty. *Psychological Review*, *102*, 269–283.
Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–314.
Tversky, A. & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, *59*, 251–278.
Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics*, *106*, 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. & Koehler, D.J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547–567.
Tversky, A. & Wakker, P.P. (1995). Risk attitudes and decision weights. *Econometrica*, *63*, 1255–1280.
Tversky, A., Sattath, S. & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*, 371–384.
von Winterfeldt, D. & Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge: Cambridge University Press.
Wakker, P.P. (2001). Testing and characterizing properties of nonaddictive measures through violations of the sure-thing principle. *Econometrica*, *69*, 1039–1059.
Wakker, P.P. (in press). On the composition of risk preference and belief. *Psychological Review*.
Wakker, P.P. & Deneffe, D. (1996). Eliciting von Neumann–Morgenstern utilities when probabilities are distorted or unknown. *Management Science*, *42*, 1676–1690.
Wu, G. & Gonazalez, R. (1996). Curvature of the probability weighting function, *42*, 1676–1690.

Wu, G. & Gonzalez, R. (1998). Common consequence effects in decision making under risk. *Journal of Risk and Uncertainty*, *16*, 113–135.

Wu, G. & Gonzalez, R. (1999a). Dominance violations and event splitting in decision under uncertainty. Unpublished paper, University of Chicago.

Wu, G. & Gonzalez, R. (1999b). Nonlinear decision weights in choice under uncertainty. *Management Science*, *45*, 74–85.

Yates, J. & Zukowski, L.S. (1976). Characterization of ambiguity in decision making. *Psychological Science*, *21*, 19–25.