Tversky, A., Sattah, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review, 95,* 371–384.
Wilson, T. D., Lisle, D. J., Schooler, J. W., Hodges, S. D., Klaaren, K. J., & LaFleur, S. J. (1993). Introspecting about reasons can reduce post-choice satisfaction. *Personality and Social Psychology Bulletin, 19,* 331–339.
Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60,* 181–192.

11

A Rational Decision Theory

There is nothing more profitable . . . than to take good counsel with [oneself]; for even if the event turns out contrary to one's hopes, still one's decision was right.

—Herodotus

11.1 Formally Defining Rationality

We have made frequent references to rational choice processes. Now it's time to describe the reigning rational (normative) decision theory. Some experts have defined *rationality* in terms of compatibility between choice and value: Rational behavior is behavior that maximizes the value of consequences. But, as should be clear by now, the question of what constitutes a value is not easily answered, and we think that rationality of choice is a matter of the process of choosing, not of what is chosen. Nevertheless, some very important research in decision theory is concerned with the relationship between decisions and the values of the decision makers. This is the work of John von Neumann and Oskar Morgenstern (1947), in particular their classic analysis described in *Theory of Games and Economic Behavior*. Their *expected utility theory* is the most general and popular description of rational choice in the mathematical and behavioral sciences. We will introduce this framework and relate it to the psychology of decision making in this chapter.

We have mentioned, several times, that we (and most psychologists) believe rational theories are at best an approximate description of how humans really behave. While most people seem to realize that their actual behavior and rational standards diverge, they still want to make good decisions. They want to avoid contradictions in their reasoning and in their behavior; they usually want to behave consistently with the principles of rationality laid down in the expected utility theory we describe in this chapter. Just as with probability theory, as a species we are not endowed with a natural, intuitive sense of these principles. This lack of a clear intuition provides a good reason for an examination of the von Neumann and Morgenstern theory; it doesn't come naturally, so we need to study it to understand its implications for our behavior. To that end, we will also present a perspective on how expected utility theory can be used for *improving* the quality of decision making.

Von Neumann and Morgenstern's (1947) work was purely mathematical. They demonstrated that if a decision maker's choices follow certain (rational) rules ("the axioms"), it is possible to derive *utilities*—real numbers representing personal values—such that one alternative with probabilistic consequences is preferred to another if and only if its *expected utility* is greater than that of the other alternative. Let us break their argument up into a series of steps:

1. It begins by assuming that a decision maker's choices among alternatives with probabilistic consequences "satisfy the axioms" defining rational choice.
2. Then it is possible to associate a real number with each consequence that can be termed the *utility* of that consequence for that decision maker.
3. The *expected utility* of a particular alternative is the expectation of these numbers—that is, the sum of the numbers associated with each possible consequence weighted by the probability that each consequence will occur.
4. The conclusion is that a decision maker will prefer outcome X to outcome Y if and only if the expected utility (number) associated with X is greater than that associated with Y.

The axiomatic system achieves several important goals. First, it spells out succinctly and precisely a list of principles of rational decision making. Of course, even at the normative, philosophical level these principles are a hypothesis about the essence of rational decision making. Other philosophers and mathematicians have proposed alternative systems to prescribe rational decisions, although there is no question that the von Neumann and Morgenstern system is the current champion. Second, if the axioms are satisfied, then a scale of utilities can be constructed in which the real numbers represent the values of consequences in an orderly manner. In a moment, we will develop an analogy to scales for physical weight; it should be obvious how useful to progress in science and practical applications it is to have such numbers. (Imagine modern physics, chemistry, and engineering without

numerical scales.) Third, although it is not specified in detail in the axioms, they provide a method to scale utilities, using human preferences for various outcomes as the inputs.

Many decision theorists who concentrate on the relationship between values and action define rationality as making choices that are consistent with these axioms, and they further hypothesize that real human choices are also described by the axioms. A rational decision maker would then be one who prefers alternative X to alternative Y whenever the expected utility of X is greater than that of Y. Of course, the system itself does not require a decision maker's choices to satisfy the axioms; the axioms are a hypothesis about ideal rational choices. In fact, much of the psychology of judgment and choice, already reviewed in the first 10 chapters of this book, implies that people do *not* satisfy the axioms in many decision-making contexts.

There is also nothing in von Neumann and Morgenstern's (1947) expected utility theory that states that the person making the decision has any insight into his or her utilities. The utilities are purely mathematical entities, and their existence is defined by the axioms—just as the lines and vertices of triangles we study in geometry are mathematical entities defined in terms of the axioms of that system. Nevertheless, just as we identify the abstract ideas of points and lines in geometry with the points and lines in the physical world—or on a piece of paper or pictured in our minds—these utilities are often identified with the personal values of the decision maker. Because humans are sentient, active agents (unlike the physical material to which geometry is relevant), there is a special confusion that arises between analytic and synthetic interpretations of the axioms. When expected utility theory is applied *analytically* (usually by economists), actual choices are interpreted as revealing preferences, and these revealed preferences are interpreted as implying utilities. The application has a *postdictive* flavor; as the psychologist Lola Lopes (1994) puts it, "In the modern [analytic] view, utility does not precede and cause preferences; it is instead merely a convenient fiction that can be used by the practitioner to summarize the preferences of those who, by choice or chance, follow the dictates of the von Neumann and Morgenstern axiom system" (p. 286).

In contrast, the theory can be applied *synthetically*. A person is asked first to make judgments about his or her utilities and probabilities, and then those judgments can be combined according to the axioms to predict that person's decisions. For most of us, this sequence makes the most sense: When we make a decision, we usually first try to figure out what we want and how to get it, and only then do we decide what action to take and what choice to make—first our goals and values are determined, then choices and actions. The analytic sequence—first we observe what we choose and then we infer what we must have wanted and expected—seems backward. However, there are some exceptions, as in psycho what our behavior implies about its precursors in desire and belief. And we must remember that the analytic interpretation of decisions is equally valid, and more popular than the synthetic interpretation, among the experts in economics and mathematics who are the primary users of the theory.

Nevertheless, as Tversky and Kahneman (1974) point out in a classic paper, even the probabilities are usually estimated first—and then "used" to make decisions. They write the following:

> It should perhaps be noted that, while subjective probabilities can sometimes be inferred from preferences among bets, they are normally not formed in this fashion. A person bets on team A rather than on team B because he believes that team A is more likely to win; he does not infer this belief from his betting preferences. Thus, in reality, subjective probabilities determine preferences among bets and are not derived from them, as in the axiomatic theory of rational decision. (p. 1130)

Early conceptions of utility in economics (e.g., Jeremy Bentham's [1789/1948] ideas) had a psychological quality, but modern utility theories have eliminated most of the psychology and retained only the behavioral principle that people choose what they prefer. However, there has been a major shift in the past decade with behavioral scientists like Daniel Kahneman, Colin Camerer, George Loewenstein, David Laibson, and others enriching the economic conceptualization with psychological content concerning the cognitive and emotional sources of value judgments.

When we talk about personal values, we have a far broader concept in mind than the concept of utility in the von Neumann-Morgenstern theoretical system. For example, we believe that people can verbalize some of their personal values or value systems; we do not infer these from behavior alone. Otherwise, our language system would not include such concepts as *hypocrisy,* which refers to a discrepancy between a stated value and a particular behavior. Moreover, we believe that values exist independently of both verbalization and behavior. In ordinary language, we regard values as an important existential dimension on which we can place objects, actions, and other phenomena. For example, we say, "He values freedom," as easily as we say, "He went to work yesterday." In fact, we often treat statements of value as if they were statements of fact, even though many philosophers make a very strong distinction between these two types of statements, and only after studying philosophy do most of us become confused by our own beliefs that we or others value certain objects or actions. (Some logical positivists have argued that statements that refer to values are arbitrary, or at least that such statements have no empirical referents.) Perhaps we should (another value!) be less cavalier in our everyday thinking and speaking. The research reviewed in Chapters 9 and 10 should warn us that many intuitive beliefs about our personal values are of dubious validity.

Another important characteristic of values is that they transcend particular situations. When we say we value something, we are referring to more than our behaviors, feelings, and beliefs in just one particular situation. "He values freedom," for example, refers to a general set of dispositions, actions, and beliefs, and once again, a set that a person can at least vaguely verbalize. In fact, there is a popular personality test, the Rokeach Values Inventory, that asks respondents to rank order the entries in a list of abstract value terms—equality, freedom, family security, wisdom, religious salvation—and then uses individual rankings to predict individual behaviors. For example, people who give the term *equality* a high ranking are likely also to support political policies such as school integration, affirmative action, and programs to benefit racial minority group members. Analogously, those ranking *salvation* highly are likely to be regular churchgoers. (We would speculate that the predictive power of the Rokeach test derives from the principles we discussed concerning the belief sampling model for the construction of summary values in Section 9.5. The value labels [e.g., "equality"] sample from the same pool of related, evaluatively loaded beliefs that are sampled when we are in relevant situations [e.g., have just been asked to sign a pro-affirmative action petition]; hence, the predictability stems from an overlap of memories retrieved with value-relevant cues.)

11.2. Making Theories Understandable—The Axiomatic Method

The Greek mathematician Euclid was the first we know of to summarize theories (geometry and number theory, in his case) as elegant, brief, axiomatic systems. The idea was that the essential assumptions of a theory would be extracted and written down in a precise notation, and then implications of the core theory (e.g., theorems) would be derived from those axioms. The benefits of such an approach to theoretical expression were numerous: Theorists could check for the completeness and consistency of the original theoretical statements. Scientists could focus on the essentials when applying, testing, and revising theories. And disagreements over what a theory assumed or what it implied could be resolved in a systematic and productive manner. However, axiomatization is far from universal in the sciences, and the method is even limited in mathematics. Utility theory is one of the behavioral science theories to have been axiomatized, and this has given it a great advantage in competition with less orderly formulations.

As an axiom system that leads to derivations of numerical utilities, the von Neumann-Morgenstern theory is of special interest to behavioral scientists because its conclusions have implications about decisions and values as we understand these terms in everyday language and life. Just as the conclusions of Euclidian geometry can be applied to real-world objects, we suppose the conclusions from expected utility theory describe or can be compared to human decision behavior—otherwise, they would simply be systems of rules for manipulating symbols and deriving numbers that would have little interest for most of us.

To explain the nature of a mathematical axiom system and how such a system can be related to real-world objects and phenomena, we will start with a system that is simpler and more concrete than von Neumann and Morgenstern's, but which has an analogous structure. Specifically, let us consider the weights of physical objects. Such weights are *positive real numbers;* they can be added together, as when a 137-gram weight and a 786-gram weight are put together on a scale to yield 923 grams. Such real numbers have important properties, eight of which are elaborated here.

Property 1. Comparability: Given any two positive real numbers, one is larger than the other or they are equal. That may be expressed algebraically by letting x and y stand for the numbers. Then $x > y$, $y > x$, or $x = y$. To avoid expressing all of the following properties in terms of both inequality and equality, we will usually use the "weak" form of comparability: "greater than or equal to," symbolized $\geq$. Thus, we can express comparability as meaning that for any two real numbers x and y, $x \geq y$, $y \geq x$, or both (in which case they are equal).

Property 2. Ordering: The relationship "greater than or equal to" determines the transitive ordering of the numbers; that is, if $x \geq y$ and $y \geq z$, then $x \geq z$.

Property 3. Additive closure: When we add two positive numbers, we get a third positive number; that is, if x and y are positive numbers, $z = x + y$ is a positive real number.

Property 4. Addition is associative: The order in which we add numbers is unimportant; that is, $x + (y + z) = (x + y) + z$.

Property 5. Addition is symmetric: The order in which two numbers are added is unimportant; that is, $x + y = y + x$).

Property 6. Cancellation: When a third number is added to each of two numbers, the order of the two sums is the same as the order of the two original numbers; that is, $x + z \geq y + z$ if and only if $x \geq y$.

The next two properties are more mathematical in motivation. The *Archimedean* property (although it is credited to Eudoxus, circa 408–355 BCE) asserts that no positive real number is infinitely larger than any other; that is, no matter how much smaller one number is than a second number, there is some multiple of the first number that is larger than the second number.

Property 7. The Archimedean property: Given any two numbers, there always exists an integer-value multiple of one that is larger than the other; that is, if $x \geq y$, then there exists an n such that $ny > x$. Here, ny simply refers to y added to itself n times; this axiom does not involve the general concept of multiplication, since multiplication is unnecessary when we combine integers. (Note that the Archimedean property implies that there is no largest or smallest positive real number: Given any two numbers x and y with $x \geq y$, x cannot be the largest number because there exists an n such that $ny > x$. Similarly, y cannot be the smallest, because $y > x/n$ with that same n.)

Property 8. Solvability: If $x \geq y$, there exists a z such that $x < y + z$.

The German physicist and mathematician Hölder recognized that the way in which objects behave on a pan balance corresponds perfectly to the eight axioms of this system, where $x\ R\ y$ indicates that object x outbalances object y (R can be thought of as a physical interpretation of the mathematical $>$ relationship), and the operation O corresponds to placing two objects together on the same pan (*concatenating* them). Readers should confirm for themselves that the behavior of objects on a pan balance satisfies these eight properties restated as abstract axioms where R refers to the tilt of the balance and O refers to placing objects in the same pan. (This correspondence is *conceptual;* any particular pan balance may not be large enough to hold all objects that have weight or may be subject to errors in its actual operation.) Thus, Hölder demonstrated mathematically the correspondence between his axiom system and the positive real numbers, and noted the empirical correspondence between the axiom system and the behavior of objects on pan balances. The result is that the behavior of objects on pan balances may be used to assign them real-number measures, called *weights.*

Then, in 1901, Hölder demonstrated something quite profound. Using knowledge from a branch of mathematics called *measurement theory,* he showed that if a system has these eight properties ("the axioms are satisfied"), then real numbers can be associated with the elements of the system, and these real numbers are unique except for multiplication by a positive constant. (A measurement scale with these properties is technically called a *ratio scale.*) That is, he restated these eight properties in terms of axioms in which an abstract relationship R replaced the $\geq$ and an abstract operation O replaced addition. He subsequently demonstrated that if the elements of the system related by R and combined by O satisfy these eight axioms, then it is possible to associate a positive real number with each, such that (1) the real numbers associated with $x \geq y$ if and only if $x\ R\ y$, and (2) $z = x + y$ whenever

$z = x \bigcirc y$. Moreover, any two sets of real numbers associated in this manner have the relationship that one set is a positive multiple of the other. For example, the number of kilograms assigned to an object is 1,000 times the number of grams. These numbers are termed *measures;* the measure associated with the entity x is often symbolized $m(x)$. Just as a 1 is the unit of measurement in the real numbers, the *standard gram* or *standard ounce* is the unit of measurement for weights. Finally, the fact that we measure weights in grams or ounces—which are related multiplicatively—corresponds to the conclusion that two sets of measures assigned to the entities satisfying the axiom system are positive multiples of each other.

11.3 Defining Rationality: Expected Utility Theory

The axiomatic method, stating the essentials of a theory as an elegantly simple set of postulates that include all and only the necessary definitions and assumptions from which the entire theory can be derived, is a brilliant intellectual invention. Although the method is far from universal in the sciences, we believe it is reasonable to propose the axiomatic formulation as an eventual goal for the expression of all significant theories in any scientific discipline. Even where the theory is precisely stated in equations, a computer program simulation, or words, a summary of the essential principles is absolutely necessary for comprehension and evaluation of any theory (see, for example, Hastie & Stasser, 2000). There is not a unique axiomatization for each theory; usually different but logically equivalent axiom systems can be stated. Our choice of the particular eight properties of numbers, and hence the translation into the eight axioms of weight, is based on our judgment of which axioms readers will find easiest to understand. Other authors cite different systems, many of which are more elegant and may be easier for an expert to work with. In the case of the von Neumann and Morgenstern utility theory, there are many axiomatizations, and we have chosen the one that we believe is most comprehensible to an intelligent reader who is not an expert in mathematical logic.

The basic *entities* of the von Neumann-Morgenstern system can be conceptualized as *alternatives* to be evaluated or chosen between, consisting of probabilistic consequences—often referred to as *gambles.* The basic relationship is one of *preference,* which induces an order on the alternatives; we will settle for the *weak ordering,* $\geq$, which might be expressed as "is indifferent or preferred to" behaviorally. (Note that this expression is rough; a more precise interpretation of $\geq$ is "is not unpreferred to," because both $A \geq B$ and $B \geq A$ are possible, in which case $A \sim B$; however, that usage is quite awkward. To be technically precise, we should distinguish between strong preference ($>$) and weak preference ($\geq$); the reader wishing such precision can translate by considering weak and strong preference separately.)

The basic *operation* for combining the alternatives (analogous to placing more than one object on a balance pan—the general term is *concatenation*) may be conceptualized as a *probability mixture* of alternatives. Thus, if A and B are alternatives, ApB refers to receiving alternative A with probability p and alternative B with probability $(1 - p)$. Note that the probability of receiving B is implicit, and we will only consider binary, two-alternative gambles, so given the probability of A (e.g., p), the probability of B will be 1 minus that value $(1 - p)$ or the *complement.*

To illustrate some relationships that will be prescribed by the axioms, consider the evaluation of a complex, multistage gamble. The following results will be explicated when we discuss the interpretation of individual axioms (especially Axiom 3, "Closure," and Axiom 4, "Distribution of Probabilities"). Because the alternatives specify the consequences with particular probabilities, the probability mixture of alternatives is synonymous with a probability mixture of the consequences; that is, if alternative A consists of consequence x with probability r and consequence y with probability $(1 - r)$, whereas alternative B consists of consequence z with probability s and alternative w with probability $(1 - s)$, then ApB consists of consequence x with probability rp, consequence y with probability $(1 - r)p$, consequence z with probability $s(1 - p)$, and consequence w with probability $(1 - s)(1 - p)$. An alternative with a single consequence is conceptualized as one in which the consequence occurs with a probability of 1.

We find it helpful to represent alternatives, consequences, and probability mixtures in the decision tree diagram format, as in Figure 11.1. These diagrams are useful when theoretical gambles are to be compared (e.g., to grasp the implications of the axioms or the structure of experimental stimulus gambles), and especially when the system is to be applied to analyze actual decisions.

What von Neumann and Morgenstern (1947) proved is that when their axioms are satisfied, a numerical measure can be associated with each consequence—termed a *utility* of that consequence (analogous to the numerical weight of a physical object)—and that the alternatives themselves can be ordered according to their expected utility. In other words, the basic result is that a preference between the alternatives can be represented by an ordering of their expected utilities. (Because a single consequence can be conceptualized as an alternative in which that consequence occurs with the probability of 1, and vice versa, the axioms can be stated in terms of either consequences or alternatives. We chose to present the axioms in

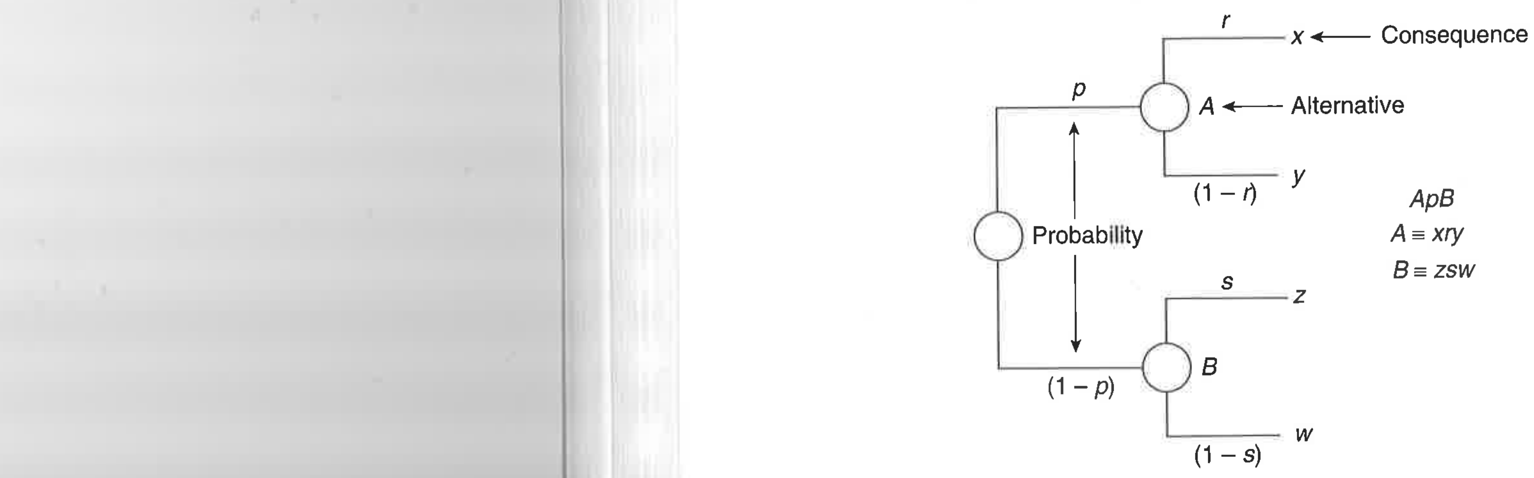

Figure 11.1 Example decision tree representation of a gamble that might be abbreviated in the von Neumann and Morgenstern axioms

terms of alternatives because we think they will be easier to understand in that form.)

We can summarize the analogy between the von Neumann and Morgenstern system and Hölder's axiomatization of physical weight as follows: Alternatives in a choice set are analogous physical objects to be weighed; weak preference is analogous to the "weighs the same as or more than" result from a pan balance test; and the "probability mixture" concatenation operation (mixing two alternatives together in a binary gamble) corresponds to placing more than one object on a pan at the same time. In both systems, if the axioms are satisfied, the result is a real-number scale, of utility or of physical weight.

Here, at last, are the von Neumann and Morgenstern axioms for expected utility theory:

Axiom 1. Comparability: If A and B are in the alternative set S, then either $A \gtrsim B$ or $B \gtrsim A$ or both, in which case $A \sim B$.

Axiom 2. Transitivity: If $A \gtrsim B$ and $B \gtrsim C$, then $A \gtrsim C$.

Axiom 3. Closure: If A and B are in alternative set S, then ApB is as well.

Axiom 4. Distribution of probabilities across alternatives: If A and B are in S, then $[(ApB)qB] \sim (ApqB)$.

Axiom 5. Independence: If A, B, and C are in S, $A \gtrsim B$ if and only if $(ApC) \gtrsim (BpC)$.

Axiom 6. Consistency: For all A and B in S, $A \geq B$ if and only if $A \geq (ApB) \geq B$.

Axiom 7. Solvability: For all A, B, and C in S, if $A \gtrsim B \gtrsim C$, then there exists a probability p such that $B \sim (ApC)$. (This axiom is crucial to the construction of the utility scale.)

If real numbers are substituted for the alternatives and probabilities for the p's and q's, then it is clear that these axioms are satisfied whenever the number associated with an alternative is equal to its expectation. What von Neumann and Morgenstern (1947) did was prove the converse: If these axioms are satisfied, then it is possible to construct a measure for each alternative equal to its expectation in such a way that the order of the alternatives corresponds to the order of the expectations. Moreover, the origin and units of these measures are arbitrary (as in the familiar scale of temperature). The number associated with an alternative is termed its *expected utility;* that number associated with a consequence, which is equivalent to an alternative that has that consequence with probability 1, is the utility of that consequence. Because only the origin and the unit of measurement are arbitrary in such utility assignments, any different assignments are related in a linear manner, meaning any two scales of utility across a set of alternatives will plot as a straight line in x-y graphic coordinates. Technically, this is called an *interval scale* because both the units of measurement and the "zero point" are arbitrary. (Remember that we called the weight scale a *ratio scale,* because the zero point was *not* arbitrary, although the units [e.g., grams, ounces] were.)

The arbitrary origin and unit of measurement allow us to use the solvability axiom to determine the utility of a third alternative whenever the utilities of two others are known. Suppose, for example, $A \gtrsim B \gtrsim C$. We can allow the utility of A to equal 100 and the utility of C to equal 0. Now, according to the solvability axiom, there exists a probability p such that the utility of B is equal to the utility of ApC that is simply p times the utility of A plus $(1 - p)$ times the utility of C, which is $p100 + (1 - p)0$. Thus, as promised, the solvability axiom is crucial in determining the actual numerical values of these utilities. Because all possible scales and utilities are linear functions of each other, we can assign 100 as the utility of the most preferred alternative in each set S, and 0 as the utility of the least preferred alternative, and then solve for the utilities of all the remaining alternatives, locating their utilities in the 0–100 interval. Notice that the probability scale in the probability mixture operation is the means to scale the utilities. (The laws and scale of probability will be preserved in the decision maker's preferences *if* "the axioms are satisfied"—an assumption that may be hard to satisfy behaviorally, if the empirical results reported in Chapters 1–10 are valid.)

Von Neumann and Morgenstern's system is conceptually beautiful. At the risk of being repetitive, we state again that the *utilities* derived analytically from these axioms do not necessarily correspond to our intuitive or verbal notions of personal value, any more than the measures of weight derived from behavior of objects in pan balances necessarily correspond to our intuitive notions of weight. Nevertheless, just as a concept of weight that did not relate to our intuitions about which objects are heavier than which others would be a very strange notion indeed, the concept of utility is meant to have a relationship to subjective value. In fact, it is because the utilities derived according to the von Neumann and Morgenstern system *do* bear some relationship to our notions of personal value that they are of interest to psychologists. And this is the justification for our attempt to explain how they can be used to *improve* our decision-making capabilities at the end of this chapter. Most people, perhaps after some thought, acknowledge that each axiom individually seems to be acceptable as part of a general definition of rationality, and even as a prescription for how they wish they made their choices. Let's discuss each axiom in more detail.

Axiom 1. Comparability: If A and B are in the alternative set S, then either $A \geq B$ or $B \geq A$, or both, in which case $A \sim B$.

Axiom 1 states that when faced with two alternatives, the decision maker should have at least a weak preference. The strongest rationale for this axiom is the fact that a decision maker faced with alternatives must choose one of them. But it also equates inability to make such a choice with indifference. Is someone who maintains that he or she cannot choose between two alternatives necessarily indifferent? Consider, for example, the choice discussed in Chapter 10 of the professor trying to decide what job to take. If she were to conclude that she could not make a choice, would that really mean that she is indifferent, that she does not care? Jay Kadane, Mark Schervish, and Teddy Seidenfeld (1999), for example, maintain that not having a preference is *not* equivalent to being indifferent. And in some circumstances where "protected values" are perceived to be at stake, people refuse to make choices—which is, of course, a choice, too. Jonathan Baron and Mark Spranca (1997) cite situations in which many people refuse to choose. For example, when considering personal or policy alternatives where trade-offs between lives and money are demanded, many citizens appear to avert their gaze and "choose not to choose." But is it reasonable to say that these non-choices express a true indifference between 55 mph speed limits and inexpensive automobiles versus the deaths of approximately 50,000 fellow citizens from traffic accidents?

Apples and oranges are, however, both fruit, and if one must choose a fruit from a dish of apples and oranges, it will be either an apple or an orange. Could not the choice itself define the preference, analytically? Economists refer to such a choice as a *revealed preference*, and assume utility theory to infer that, for example, our highway safety preferences imply a dollar value of approximately $3,000,000 per human life. Moreover, isn't it true that when people say they have no idea why they made a choice, subsequent questioning often reveals that there really *is* a preference involved? For example, if the professor in Chapter 10 maintained that she really was incapable of choosing between jobs but "just happened to pick" one of them in order to be near (or away from) relatives, would not proximity to relatives be an important consideration in her choice? Perhaps she simply would be unaware of this factor at the time she made the choice—or perhaps embarrassed to discuss it, because she might not consider it a good reason for choosing one job over another. Our own position is that people really do have preferences, except in rare instances such as predicting the outcome of a coin toss, in which case they are truly indifferent. We do not, however, accept the revealed preference position—that the preference is inherent in the choice—for the reasons outlined in the previous chapters. Specifically, choice may be truly irrational, and hence, contradictory. Thus, it follows that there may be a discrepancy between choice in a particular situation and the preferences of the individual making the choice.

While revealed preference can be rejected on the basis of the cognitive difficulty of choice, the most common reason for rejecting the apparent evidence is that people sometimes do things they do not *want* to do; that is, they choose alternatives they do not prefer. For example, the psychologist and philosopher William James (1842–1910) noted that people with toothaches often prod the painful area of their mouth with their tongue, although they clearly prefer lack of pain to the pain that results from prodding.

The counterargument from the revealed preference theorist is that the very act of prodding the area of a toothache indicates that the individual has a greater positive value for the information gained that the tooth is still hurting than negative value for the pain experienced. Such values may appear "stupid," because toothaches tend not to go away on their own without treatment, and when a given part of our mouth is aching we can be more than reasonably sure that it will be more painful if we touch it with our tongue, *without* actually doing so. The revealed preference theorist, however, has the counterargument that *de gustibus non disputandum* ("There's no disputing matters of taste"). The fact that the sufferer prods the tooth reveals that even such redundant information is worth the pain.

Because what constitutes pleasure and pain to an individual cannot be known unambiguously, the argument that people often do what they really dislike doing is fairly ineffective against the revealed preference position. In contrast, knowing that the choices are often contradictory for *cognitive* reasons undermines this position.

Axiom 2. Transitivity: If $A \geq B$ and $B \geq C$, then $A \geq C$.

The primary justification for Axiom 2 is that individuals who violate it can be turned into "money pumps." Suppose that John Dolt preferred alternative A to alternative B, alternative B to alternative C, and yet C to A. Assume, further, that he is not indifferent in his choice between any of these alternatives. Consequently, he should be willing to *pay something* to trade a less preferred alternative for a more preferred one. Now suppose John is given alternative C *as a gift*. Because he prefers alternative B to C, he should be willing to pay something to have B instead. Subsequently, John should be willing to pay something to have alternative A substituted for B, and finally to pay for the substitution of C for A. Then John will have paid three times for the privilege of ending up with the alternative he was given in the first place. By repeating this cycle indefinitely, John (hypothetically, anyway) would pay a lot of money to get nowhere.

The response to the money pump argument is that an individual with intransitive preferences would simply refuse to play that game. Choices are, after all, not made repeatedly, but in a particular context. A choice between two alternatives does not have to be one by which the individual is bound for all time and in all circumstances. One noted economist is quoted as saying that in a particular decision-making situation, most people will "satisfy their preferences and let the axioms satisfy themselves" (Paul Samuelson, quoted by Daniel Ellsberg, 1961). For example, consider the hiring of a new secretary. Suppose that the employer has three criteria for making a job offer: (1) clerical skills, (2) organizational ability, and (3) willingness to run errands and do other jobs not specifically in the position description. Suppose that the rank of three prospective secretaries (A, B, and C) on clerical skills is A, B, C; on organizational ability is B, C, A; and on willingness is C, A, B. Thus applicant A is superior to applicant B on two of these three dimensions (clerical skill and willingness), B is superior to C on two (clerical skill and organization), and C is superior to A on two (organization and willingness). A decision based on the principle that one applicant is preferred to another whenever that applicant is superior on two of three dimensions results in intransitivity. (This is the qualitative additive difference or *voting rule* choice strategy discussed in Chapter 10.) What will happen is that the *order* in which the applicants are considered will be crucial, with the applicant considered last being the one chosen.

Is this consequence necessarily a bad one? Even though the employer could in principle become a money pump, no one is going to make her one—by giving her one of the secretaries and then demanding payment for subsequent substitutions. But we suspect that many of our most frustrating decision experiences occur when we encounter alternatives that are important, and where each alternative has some good and some bad attributes, forcing us to consider compensatory trade-offs, and sometimes inducing intransitivity-producing strategies: "Okay, I'm going to take the high-paying job; but no, I don't want to give up the much more flexible vacation options in the second job; but wait, I don't want to live in the Midwest; but the chance for advancement is much better in the third job . . ." This sometimes produces great personal discomfort and leads to an inability to make a choice.

We believe that choices *should* be transitive. This idea is part of a general argument that choice is superior when it is made as if the decision maker *were* bound by that choice in a broader context and across time. This argument is from Immanuel Kant (1724–1804), who proposed that individuals should make choices as if they were formulating a *policy* for all people at all times. There is empirical evidence showing that when a criterion is available according to which we can decide whether choices are good or bad, choices made in accord with Kant's principle are in fact superior to those made in the narrower context of considering only the options immediately available.

Axiom 3. Closure: If A and B are in alternative set S, then ApB is as well.

Axiom 3 simply requires that the decision maker be capable of conceptualizing a probability mixture of alternatives as itself an alternative. If people were incapable of doing so, there would be little point in theorizing about decision making.

Axiom 4. Distribution of probabilities across alternatives: If A and B are in S, then $[(ApB)qB] \sim (ApqB)$.

Basically, Axiom 4 requires that people follow the principles of probability theory (see Appendix). This axiom is illustrated in the decision trees in Figures 11.2 and 11.3; the two-stage gamble on the left must be treated as equivalent to the one-stage gamble on the right to satisfy the axiom.

Of course, people may violate it without disputing it; for example, it is violated by the person who reacts differently to a consequence of receiving \$45 with probability .20 than to a two-stage consequence in which the person receives nothing with probability .75 in the first stage and then receives \$45 with probability .80 if the second is reached. (Since $[1.00 - .75] \times .80 = .20$, the distribution axiom requires that the two lotteries be treated as identical.)

Figure 11.2 Decision tree representation of the two abstract lotteries mentioned in Axiom 4

Figure 11.3 Decision tree representation of a concrete pair of gambles illustrating an equivalence implied by Axiom 4

Von Neumann and Morgenstern (1947) discuss probabilities as if they were objective. (Their *p*'s and *q*'s are supposed to be probabilities measured on an absolute scale, although the concept of accuracy or objectivity in probabilities is complex and controversial.) There are several decision theories in which this axiom is relaxed—most notably Ward Edwards's early proposal to shift from objective to subjective probabilities, creating an alternative rational decision theory. It is certainly possible to have a set of nonobjective probabilities that are internally coherent and consistent with the rules of probability theory

(discussed at length in Chapter 9 and the Appendix). However, if the decision maker attempts to deal with future uncertainty by making probability assessments, these must be made according to rules of probability theory; otherwise, contradictory choices can result. In Chapter 12, we will explore prospect theory (introduced in Chapter 9), an axiomatic *non-expected utility theory* in which nonobjective and incoherent decision weights replace probabilities. Prospect theory is similar in overall structure to the von Neumann and Morgenstern theory, but it is meant to be descriptive of human decision behavior and *not* a model of rational choice.

Axiom 5. Independence: If *A*, *B*, and *C* are in *S*, $A \gtrsim B$ if and only if $(ApC) \gtrsim (BpC)$.

Axiom 5 is crucial. In fact, many decision theorists have investigated at length the effects of violating it, or of omitting it from a set of rules governing choice. At first reading, it appears innocuous: If one alternative is preferred to another, shouldn't that preference remain even though, with some specified probability, the decision maker receives neither, but a third instead? That is all this axiom states. Figure 11.4 summarizes the axiom graphically.

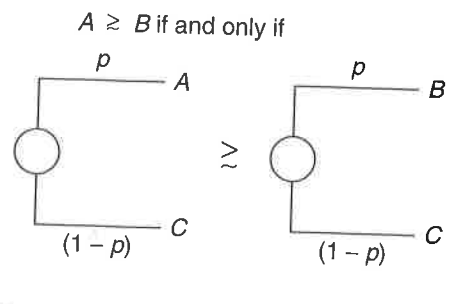

Figure 11.4 Decision tree representation of the situation described by Axiom 5

Warning: Some students are confused because they misread the axiom to refer to a situation as one of *joint receipt*—the chooser *receives both* A and C or *receives both* B and C—but the correct situation is a *probability mixture* of receiving A *or* B compared with a *probability mixture* of receiving B *or* C, where *or* is exclusive—meaning *not* both. If the situation was joint receipt, then the axiom wouldn't make much sense: Of course, we might have little desire to receive a right shoe *or* a left shoe, but to receive both a right shoe *and* a left shoe might be very attractive, depending on the shoes, of course.

Consider a pseudocertainty effect: Most people prefer a .20 probability of receiving $45 to a .25 probability of receiving $30 (Panel 1 in Figure 11.5), yet simultaneously they prefer $30 for sure to an .80 probability of receiving $45 (Panel 2). Now let *A* be the alternative of receiving $30 for sure and *B* be that of receiving $45 with the probability of .80. *A* is preferred to *B* (as in Panel 2). Let C be the alternative of receiving nothing. Now let *p* equal .25. Then (*A* .25 *C*)

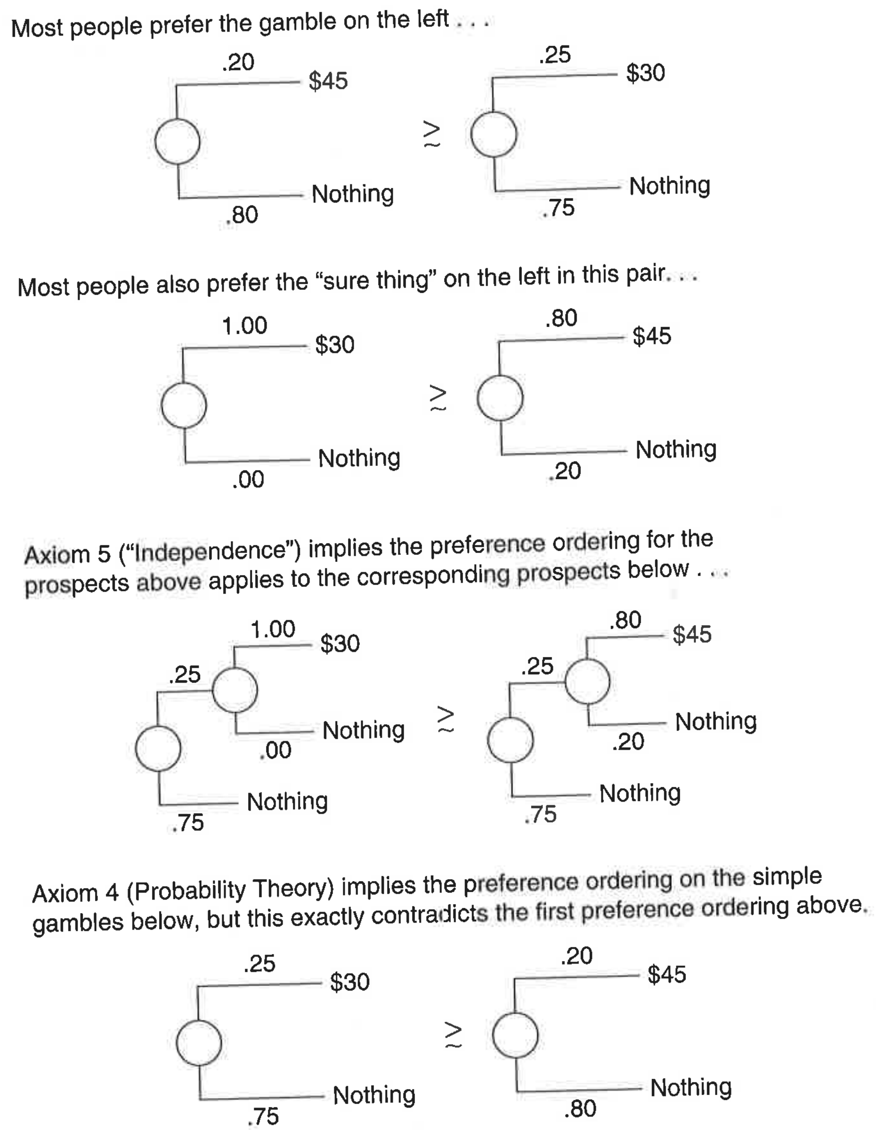

Figure 11.5 Illustration of the pseudocertainty effect that violates Axioms 4 and 5

is an alternative consisting of receiving $30 for sure with probability .25 versus receiving nothing with probability .75 (left-hand side of Panel 3), which—by the distribution of probabilities axiom—is just the alternative of receiving $30 with probability .25 (left-hand side of Panel 4). In contrast, (*B* .25 *C*) is the alternative consisting of a .25 probability of receiving $45 with probability

.80 and a .75 probability of receiving nothing (right-hand side of Panel 3)—that is, receiving $45 with probability $.80 \times .25 = .20$ (right-hand side of Panel 4). Thus, the typical preferences summarized in Panel 1 imply the opposite ordering of that summarized in Panel 2. Therefore, anyone who has the Panel 1 and Panel 2 preferences (most people do) both exhibits the pseudocertainty effect and violates the independence axiom.

The pseudocertainty effect describes choices that are influenced by the way the consequences are framed, rather than solely by the consequences themselves. Is such irrationality the only reason for violating the independence axiom? There is another reason. Axiom 5 implies that the decision maker cannot be affected by the *skewness* of the consequences, which can be conceptualized as a probability distribution over personal values. Figure 11.6 shows the skewed distributions of two different alternatives. Both distributions have the same average, hence the same expected personal value, which

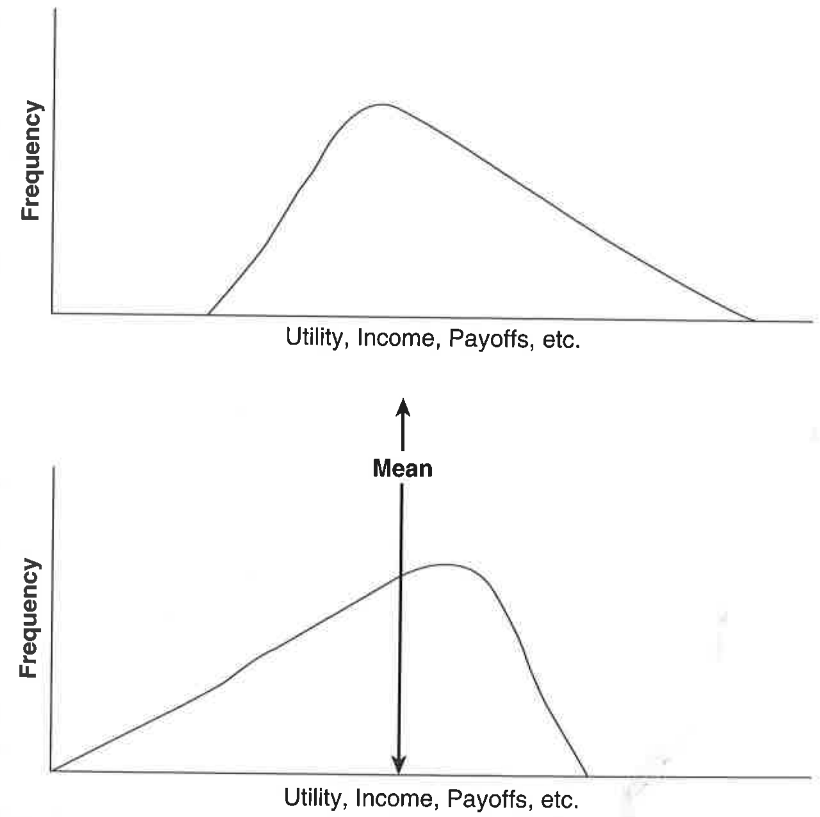

Figure 11.6 Two skewed distributions with the same average value and variance

is a criterion of choice implied by the axioms. These distributions also have the same variance. (For a description of the mean and variance of a probability distribution see a good introductory statistics text, e.g., *Statistics* by Freedman, Pisani, Purves, & Adhikari, 1991.)

If the distributions in Figure 11.6 were those of wealth in a society, most people would have a definite preference for the upper distribution; its positive skewness means that income can be increased from any point—an incentive for productive work—and incidentally it is the distribution that describes wealth in industrialized societies. Moreover, those people lowest in the distribution are not as distant from the average as in the lower distribution, where a large number of people are already earning a maximal amount of money, and there is a tail of people in the *negatively skewed* part of this distribution who are quite far below the average income. If we have such concerns about the distribution outcomes in society, why not about the consequences for choosing alternatives in our own lives? In fact, many of us do not like alternatives with large negative skews. Note also that popular lotteries, gambling devices, and competitive tournaments generally have positive skews (i.e., small probabilities of winning a lot). There are substantial individual differences in preferences for distributions of multi-outcome gambles or lotteries; in experimental studies in which money lotteries are evaluated, positive skewed lotteries (like the upper distribution in Figure 11.6) are the modal favorites for lotteries composed of both gains and losses (Lopes & Oden, 1999).

Figure 11.7 Decision tree representation of the relationship expressed in Axiom 6

Axiom 6. Consistency: For all A and B in S, $A \geq B$ if and only if $A \geq (ApB) \geq B$.

Axiom 6 states that if we prefer one alternative to another, then we prefer at least some chance of receiving that alternative rather than the other one (see Figure 11.7). This axiom appears indisputable.

Axiom 7. Solvability: For all A, B, and C in S, if $A \geq B \geq C$, then there exists a probability p such that $B \sim (ApC)$.

Axiom 7 is similar to the Archimedean property in Hölder's axiom system for physical weights. What it states in effect is that no alternative is so much better or worse than another that some probability mixture of alternatives on either side is not regarded as equivalent to the original alternative. Consider, for example, three alternatives, A, B, and C, with the preference order ABC. The axiom states

that there will be *some* probabilistic way of combining A and C such that the individual is indifferent to choosing B or this combination (see Figure 11.8).

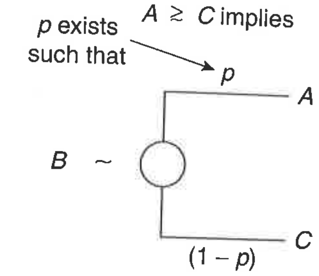

Figure 11.8 Decision tree representation of Axiom 7; the axiom implies there exists a probability *(p)* that satisfies the equivalence relationship between the "certainty equivalent" *(B)* and the gamble

Now, if A were incomparably more attractive to the decision maker than any of the other alternatives, then *any* probability of receiving A rather than C might lead to a preference for ApC over B. The same argument would hold *mutatis mutandis* (meaning "with the necessary changes") if alternative C were incomparably worse than B. The axiom states that no such alternatives exist. Well, what about eternal bliss in Heaven? Or, conversely, sudden death? Would not any alternative involving even the slightest probability of eternal bliss be preferred to some other alternative with drabber consequences—so that the individual could never be indifferent in choosing between the drab alternative and a probability mixture involving such bliss? Or do we not eschew completely those alternatives involving some probability of death? (Perhaps we should not discuss eternal bliss, because we cannot even conceptualize it.) It is clear from our behavior that we dread death and attempt to avoid it, at least for as long as we have hope that the positive aspects of life and the future outweigh the negative. Do we not, then, avoid all alternatives that involve some probability of death? The answer is, "No." Everyday life involves some risk of death, even such trivial actions as crossing a street to buy a newspaper. Sometimes this probability is more salient than at other times—as when people who fear airplane trips still fly tens of thousands of miles a year. But it is always there. Even staying in bed all day to avoid what appears to be the risk of death would involve a risk of physical deterioration, perhaps leading to death. In addition, there are clear examples of thoughtful decisions that involve a high probability of death—for instance, a decision to join an underground resistance movement during a military occupation or the deliberate choice of a high-paying, but dangerous profession like industrial deep-sea diving.

All of the axioms appear quite reasonable. In fact, if we assume comparability, we can violate only the independence axiom without becoming outright irrational. The axioms, however, have strong implications, as do other mathematical results. Believing the Pythagorean theorem, for example, we anticipate the length of a third side of a right triangle when we know the length of two sides. If physical measurement does not confirm our expectations, we conclude that the triangle is not a right triangle; we rarely or never conclude that the theorem is true—that the triangle has a right angle and our measurements are correct—but that the logic of mathematical deduction just doesn't apply. The demanding aspect of the von Neumann and Morgenstern axioms is that if we accept them, we are bound to evaluate alternatives in a choice situation in terms of their expected utility. That is, numbers exist that describe the utility of each consequence of alternative choices. (Such numbers are, once again, those associated with alternatives that have that particular consequence with a probability of 1.) These numbers can be determined by some choices, using the solvability axiom; they then require that other choices as well be made in terms of the expected utilities computed. Other characteristics describing the distribution of consequences—for example, its skewness—are irrelevant.

The solvability axiom is especially useful if we want to design a method to scale a person's utilities for outcomes. A review of measurement and scaling methods is beyond the scope of this book (see Dawes & Smith, 1985, for an introduction), but here is an example to convey the basic method. For simplicity, consider three outcomes (which could be dollars, but let's consider something less quantitative): a 1-week vacation in the city of Boulder (Colorado), Pittsburgh (Pennsylvania), or Lubbock (Texas). Suppose further that our decision maker prefers them in that ordering, Boulder over Pittsburgh, Pittsburgh over Lubbock ($A \succsim B \succsim C$ as required by the solvability axiom). Now, where in the interval between Boulder and Lubbock is the decision maker's utility for Pittsburgh? Relying heavily on "solvability," we can assign the most favored and least favored options the values of 1.0 and 0.0 (or, for that matter, because the scale origin and units are arbitrary, 100 and 0) and present the decision maker with a series of gambles, mixing a trip to Boulder with a trip to Lubbock, until we find a gamble that is judged indifferent compared with the trip to Pittsburgh (the right-hand gamble in Figure 11.8). Then, if the axioms are satisfied we can use the probability mixture number from the gamble as a scale value for Pittsburgh's utility. For example, if a decision maker is indifferent between 1 week in Pittsburgh for sure and a .80 chance of ending up in Boulder mixed with a .20 chance of spending the week in Lubbock, we would scale the utility value for Pittsburgh at 80 on a 0–100 scale with Boulder and Lubbock as endpoints. This method can be generalized to scale any number of outcomes on an interval utility scale (and this method of eliciting *preference probabilities* is frequently used in applied decision analysis).

There are many studies in which utility functions are scaled, especially for the utilities of money. The forms of these functions are often used to interpret

and even to predict the behavior of people from whom they've been derived. For example, a concave curve (negatively accelerating, with diminishing marginal returns) is sometimes interpreted as implying that the person who exhibits it is risk-averse in the domain of such a curve, while the reverse, convex curve is interpreted as implying a risk-seeking attitude. According to the theory, such curves summarize analytically and predict synthetically a person's behavior in choosing real-money gambles to play in the laboratory, and they are even associated with the occupational choices of business executives. Executives in financially volatile businesses are more likely to exhibit convexity in their utility curves than executives in more placid financial environments (MacCrimmon & Wehrung, 1986; see additional discussion of the interpretation of utility curves in Section 9.3).

11.4 Traditional Objections to the Axioms

The axioms were not presented as descriptions of actual behavior, but rather as conditions of *desirable* choices. Are they? After the publication of von Neumann and Morgenstern's (1947) book, several theorists suggested that the axioms placed unreasonable constraints on choice behavior and that they should *not* be satisfied. The best-known objections consisted of two paradoxes, originally stated as conceptual puzzles and later validated in experimental studies. One of these objections was raised by the Nobel laureate economist Maurice Allais and the other by the decision theorist Daniel Ellsberg (who achieved notoriety by releasing the *Pentagon Papers,* which exposed the United States government's secret objectives in the Vietnam War).

The Allais Paradox

Maurice Allais won the Nobel Prize in Economics for his argument that the expected utility principle that results from the von Neumann-Morgenstern axiom system is too restrictive. Consider, for example, the choice between alternatives A and B involving millions of dollars:

Alternative A: Receive \$1 million with probability 1.00 (i.e., for certain).

Alternative B: Receive \$2.5 million with probability .10, \$1 million with probability .89, and nothing with probability .01.

When presented with this (hypothetical) choice, most people choose alternative A. That means that if they abide by the axioms, it is possible to assign utilities to the consequences of receiving \$1 million, \$2.5 million, or nothing in such a way that the choice of A implies a higher expected utility for it than for B. Specifically,

U($1 million) > .10 U($2.5 million) + .89 U($1 million) + .01 U(nothing).

By the solvability axiom, we can set U($2.5 million) = 1.0 and U(nothing) = 0.0. The conclusion then is .11 U($1 million) > .10.

Now consider the choice between two different alternatives:

Alternative A′: $1 million with probability .11, otherwise nothing

Alternative B′: $2.5 million with probability .10, otherwise nothing

The expected utility of alternative A′ is .11, while that of alternative B′ is .10, because we have set the utility at $2.5 million equal to 1.0. Thus, the choice of A over B requires the choice of A′ over B′. Allais argued that it was nevertheless reasonable to choose A over B *and* B′ over A′, which is in fact the choice most people make when presented with this hypothetical pair of decisions. Why accept a 1/100 chance of receiving nothing when you can receive a million dollars for sure? Conversely, given that the most probable outcome in the second choice is to receive nothing at all, why not take a 1/100 risk of getting nothing in order to increase the potential payoff by a factor of 2.5 times?

The mathematician Leonard (Jimmy) Savage (1954) gave a compelling analysis of those questions. Consider his suggestion that the probabilities of the various outcomes are illustrated by randomly drawing a chip out of a bag containing 100 chips. Of these chips, 1 is black, 10 are blue, and 89 are red. Alternative A can then be conceptualized as paying out $1 million no matter which chip is drawn. In contrast, an individual choosing alternative B receives $1 million if a red chip is drawn, $2.5 million if a blue chip is drawn, and nothing if the black chip is drawn. Now it does not matter to the decision maker which alternative is chosen if a red chip is drawn, because in either case he or she receives a million dollars; hence, a choice of A over B implies that the possibility of receiving a million dollars rather than nothing if the black chip is drawn from the 11 chips that are not red is preferred to the possibility of receiving $2.5 million rather than nothing if one of the 10 blue chips is drawn. (Sometimes the choice is interpreted as a desire to avoid the potential regret of getting nothing: "Well, dear, I had a sure million dollars, but I decided to try to "maximize" my expected value and I gambled. Unfortunately, I lost. . . .") But that preference is violated if the individual also chooses B′ over A′. Once again, the outcome will be the same if a red chip is drawn, but the individual is now showing a preference for receiving $2.5 million if a blue chip is drawn and nothing if the black one is drawn over receiving $1 million if either a blue or black chip is drawn.

		Chips Drawn From a Bag of 100		
		89 Red	1 Black	10 Blue
CHOICE	A	$1 million	$1 million	$1 million
	B	$1 million	nothing	$2.5 million
CHOICE	A′	nothing	$1 million	$1 million
	B′	nothing	nothing	$2.5 million

Figure 11.9 Illustration of Savage's analysis of the Allais paradox

Savage's (1954) example is basically a restatement of the independence axiom in concrete terms; it makes this axiom quite compelling. Savage's argument is illustrated in Figure 11.9. In the figure, it is obvious that the event of drawing a red chip, with a constant .89 probability across the gambles, should be irrelevant to the choice, because its payoff is constant within each choice set pair.

The Ellsberg Paradox

Ellsberg's (1961) problem focuses attention on the nature of the uncertainty presented by different gambles, but again it is an assault on the acceptability of the independence axiom. Imagine an urn containing 90 colored balls—30 red balls and 60 black and yellow balls. You do not know the exact proportion of blacks and yellows, only that there is a total of 60. One ball will be drawn at random from the urn. Which of the following gambles would you prefer?

Alternative I: Receive $100 if the ball drawn is red, nothing otherwise.

Alternative II: Receive $100 if the ball drawn is black, nothing otherwise.

Most people choose alternative I; the obvious interpretation is that they would rather bet on a precise, known probability of winning, than on an ambiguous uncertainty.

Now consider another pair of gambles. Again, which would you prefer?

Alternative III: Receive $100 if a red or a yellow ball is drawn, nothing if a black is drawn.

Alternative IV: Receive $100 if a black or a yellow ball is drawn, nothing if a red is drawn.

Now most people choose IV, again because they prefer the well-defined risk (60/90 balls) over the ambiguous uncertainty. (The chances of winning in alternative III could range from 31/90 to 89/90.)

The "catch" with this pattern of choices is that it violates the independence axiom, again. The table in Figure 11.10 makes this clear. The payoffs resulting from drawing a yellow ball are identical in each pair, so the preference should depend only on the pattern of payoffs for the red and black draws; the yellow outcome should be irrelevant. But the red and black patterns are identical across the two pairs, implying there should be no change in preferences from the first pair to the second. However, most people do exhibit strong preferences, but for reversed red-and-black patterns between the two pairs. Another way to express the contradiction is to note that in the first pair, the chooser prefers to bet *on* red, rather than black. But in the second pair, the chooser prefers to bet *against* red, rather than black. These

BALLS IN THE URN

		30	60	
		Red	Black	Yellow
CHOICE	I	$100	nothing	nothing
	II	nothing	$100	nothing
CHOICE	III	$100	nothing	$100
	IV	nothing	$100	$100

Figure 11.10 The gambles presented in Ellsberg's paradox

choices imply that the chooser believes red is more likely than black *and* that not-red is more likely than not-black. Moreover, the chooser is acting as though *red or yellow* is less likely than *black or yellow*, but at the same time as though *red* is more likely than *black*. It is impossible to assign probabilities—numbers consistent with the principles of probability theory—to the outcomes given these choices.

11.5 The Shoulds and Dos of the System

Of course, people do not always choose in accord with von Neumann and Morgenstern's axioms, for several reasons, including the irrationalities of judgment and valuation that have been described in the previous chapters of this book. Daniel Ellsberg (1961) had an especially astute comment:

> There are those who do not violate the axioms, or say they won't; such subjects tend to apply the axioms rather than their intuition, and when in doubt to apply some form of the Principle of Insufficient Reason. Some violate the axioms cheerfully, even with gusto; others sadly but persistently, having looked into their hearts, and found conflicts with the axioms and decided, in Samuelson's phrase, to satisfy their preferences and let the axioms satisfy themselves. Still others tend, intuitively, to violate the axioms but feel guilty about it and go back into further analysis. (p. 655)

Should we make only choices satisfying the axioms of the system? Our answer to that question is a qualified "yes." The qualification is that although we should not be *bound* by the axioms, we should *consider* them when making choices. While it is difficult to determine whether a particular decision per se satisfies or violates an axiom or a set of axioms, the fact that the axioms are true if and only if choices are made according to expected utility provides a method for considering alternative decisions. For example, consider a husband and wife with children who decide to fly on separate airplanes. This indicates (according to the theory) that the couple felt that the death of both of them would be more than twice as bad as the death of either one alone. We will analyze this example within the von Neumann and Morgenstern framework.

There are three distinct consequences in this example: Both die, one dies and one lives, or neither dies. Because the assignment of two utility values is arbitrary (given that all utility scales are related to each other in a linear manner), we can arbitrarily assign the utility −1 to the consequence that both die and 0 to the consequence that neither dies. Now let p be the probability that one airplane crashes; whether we estimate this probability objectively by looking at the airline's safety statistics or on the basis of our hunch, our conclusion is that the probability that both airplanes independently crash is $p \times p$ or p^2. This is the probability that both parents will die if they fly separately. In contrast, the probability that they both will die if they are in the same airplane is simply p. Now let the utility that *exactly one* of them dies be symbolized x (which will be a *negative* number). The choice of flying separately is interpreted within the von Neumann and Morgenstern framework as

$$p(-1) < 2p(1 - p)x + p^2(-1).$$

The first term on the right half of the equation is x times the probability that one parent will survive the trip and the other will not (which is equal to the probability that the first plane will crash and the second will not *plus* the probability that the first one will not and the second one will—that is, twice the probability that only one plane will crash). The second term is the probability that both will die on independent trips multiplied by -1, the arbitrarily assigned utility of the *both die* consequence.

Dividing by p and rearranging terms yield $x > -½$. (After the p is canceled out, move the remaining $-p$ on the right side to the left side, which yields $p - 1 < 2(1 - p)x$; dividing by $(1 - p)$ yields $-1 < 2x$. The $x > -½$ means that the death of exactly one has less than half the (negative) utility of both dying.

Our advice is that the couple might be well advised to consider whether the death of one of them is less than half as bad as the death of both. In this example, that would probably not change the decision; in fact, such consideration would most likely reinforce it. What we have done in this example is to assume that people have at least *partial* appreciation of the utility as specified by the framework. There is nothing in the framework itself that requires such insight—just as there is nothing in the physics framework for measuring the weight of objects that requires that the numbers obtained should correspond to our ideas of which objects are heavier than which. Nevertheless, just as weights measured by pan balances do correspond—at least partially—to our subjective experience of these weights, so might the utilities in the system correspond to our subjective conceptions of personal value. In fact, if there were no such correspondence in either case, there would be little reason to be interested in either axiom system.

Consider another example, this one from a medical context involving the diagnosis of a renal cyst versus a tumor on the basis of X-ray evidence. In his doctoral dissertation, Dennis Fryback studied decisions at a university hospital to test whether a kidney abnormality that appeared on an X-ray could be a cyst or a tumor (summarized in Fryback & Thornbury, 1976).

The standard procedure was for patients who appeared to be suffering from a kidney disorder to be X-rayed, and if an abnormality appeared, the radiologist interpreting the X-ray made a probability judgment about whether that abnormality was a cyst or a tumor. Then the patient would be tested directly by an invasive procedure. No procedure existed at that time, however, that tested for both a cyst and a tumor. Moreover, because there was always the possibility that an X-ray abnormality was a *normal variant,* a negative result on the test for one of these two pathologies required a subsequent test for the other. The decision Fryback studied was which test to do first. This decision was important to the patient, because the nature of each of the tests is quite different.

The test for a cyst is termed *aspiration.* It consists of inserting a large needle through the patient's back to the location of the abnormality and determining if fluid can be drawn from it; if fluid can be drawn, the abnormality is a cyst. The procedure can be accomplished in a doctor's office with a local anesthetic; the risk of a blood clot is very low; the cost is not great.

The test for a tumor is termed *arteriography.* A tube is inserted into the patient's leg artery and manipulated up to the kidney, at which point a device on the end of the tube removes a sample of tissue from the suspected spot; this tissue sample is then subjected to a biopsy. At the time of Fryback's study, this procedure required 1 day of hospitalization in preparation for the operation and at least 1 day's hospitalization after the operation. The probability that a blood clot would develop is approximately 10 times as great as that with the aspiration procedure; the patient experiences considerable discomfort in the days following the operation; and it is much more costly than the aspiration test.

Fryback found that in general the aspiration test was done first if the radiologist believed that the probability was greater than 1/2 that the abnormality was a cyst rather than a tumor; otherwise, the arteriography test was conducted. He also found that the patients, doctors, and potential patients from the general public he questioned all thought that the arteriography test was at least 10 times worse than the aspiration test. (Interestingly, discomfort, lost work days, and probability of formation of a blood clot were the major determinants of this judgment; cost was considered irrelevant—perhaps given the assumption that "insurance pays"—which is why we do not specify the cost difference in the above description.) Fryback then conducted an expected utility analysis on the assumption that the disutility of the arteriography test was 10 times that of the aspiration test. For the purposes of this analysis, we can conditionalize probabilities on the assumption that the patient has *either* a cyst or a tumor, even though the requirement that the second test be given if the first is negative arises because the patient may have *neither;* that is, we can let p be the probability that the patient has a tumor *given* the abnormality is not a normal variant, and hence, $(1 - p)$ is the probability that the patient has a cyst given that the abnormality is not a normal variant. Again, we can arbitrarily set the utility of no test at all at 0; then, setting the utility of the aspiration test at -1, the utility of the arteriography test is -10, and hence the utility of doing both tests is approximately -11.

Let the probability that the radiologist interpreting the X-ray believes the problem to be that of a tumor be p. If the arteriography test is done first, the expected disutility of the entire testing procedure is

$$p(-10) + (1 - p)(-11).$$

The second term in the expression occurs because both tests are required if the abnormality is not a tumor. In contrast, the disutility of doing the aspiration test first is

$$(1-p)(-1) + p(-11).$$

The expected utility of doing the arteriography test first will be greater than that of doing the aspiration test first (i.e., the disutility will be *less negative*) whenever

$$-10p - 11(1 - p) > (1 - p) - 11p.$$

This is true if and only if

$$11p > 10 \text{ or}$$

$$p > 10/11.$$

In other words, doing the arteriography test first is better in an expected utility framework only under those conditions in which the probability of a tumor relative to that of a cyst is greater than 10/11—that is, when a tumor is 10 times more likely. Recall that the people questioned believed that the disutility of the arteriography test was *at least* 10 times that of the aspiration test; it follows that the 10/11 figure is a *lower bound.* Yet the procedure at the hospital was to test for arteriography *first* whenever the judged probability was greater than 1/2.

This example involved at least a partial equating of the utilities in the von Neumann and Morgenstern system with the personal values expressed by people when asked. Nevertheless, believing people can assess such utilities seems to be quite reasonable; in fact, when this conclusion was communicated to the hospital physicians, the procedure was changed. An interesting sidelight is that when the radiologists' probability judgments were checked against the actual frequencies with which cysts and tumors were found upon testing, these judgments turned out to be quite accurate. Simultaneously, however, the analysis indicated that they were also most usually irrelevant, because it implied that the aspiration tests should be done first, except in those rare cases in which the radiologist was at least 10 times more certain that the aberration was a tumor rather than a cyst.

The new (since 1970) field of applied decision analysis makes use of the von Neumann and Morgenstern approach in an attempt to aid decision makers in their choices (Clemen, 1996; Hammond, Keeney, & Raiffa, 1999). It is based on the premise that people do have some insight into their personal values, but that these values may not be reflected in single choices within a particular context—especially those that tend to be made automatically or according to a standard operating procedure. What the applied decision analyst does is question the decision maker at length about values and probabilities in hypothetical situations as well as the situation in which the choice of interest is to be made. After having done so, the analyst proposes an expected utility analysis that would allow the decision maker to systematize the alternatives in making subsequent choices. Such applications can have a profound effect, as when the hospital decided to change its order of tests.

Yet another example is provided by a man who owned a company in a small town and was considering automation. His family had owned the company for many years, and the factory provided employment to a substantial number of people in the community. After receiving an economic report on the probable increase in profits that would follow automation of many of the factory jobs, he was uneasy about implementing automation, but he was not sure why. He hired a consultant who worked as an applied decision analyst in such situations. After questioning the business owner at length, the consultant concluded that the owner's real utilities in running the business had very little to do with the profit he made. Instead, the owner derived great satisfaction from giving employment to so many people in the town; doing so provided him with status and a feeling of doing something important for the community. According to the consultant's analysis of expected utility, automation would be a very poor choice, one that would decrease rather than increase the utility this man derived from running his business. When the owner was presented with the results, his response was "aha!" He then understood that providing employment was exactly what he wanted to do. In fact, his reluctance to automate in the face of rather conclusive data that it would increase his profits was, he now understood, due to that desire. In addition to reinforcing his gut impression that something was amiss with the automation plan, the consultant's analysis gave him a rationale to use in explaining his refusal—both to himself and to those who might regard him as a poor businessperson for having bypassed the opportunity to increase his profits.

Such decision analysis is a form of psychotherapy because it helps people to change their behavior to be consistent with their personal values. The von Neumann and Morgenstern framework does not dictate what choice must be made, but it is an important tool for such therapy. Moreover, it can help prevent more basic irrationalities, because a decision made within this framework cannot be irrational. Consider, for example, an individual who is reluctant to abandon a sunk cost; because expected utility analysis concerns only the expectation of *future* consequences, sunk costs are not entered into the analysis. In effect, the individual who is tied to the sunk costs broadens his or her framework in such a way that the original motive—the perception of waste in abandoning the sunk costs—disappears when the person realizes that honoring sunk costs conflicts with the more important motive of behaving in an economically rational manner.

The decision analyst starts with the assumption that there are conflicts between a client's attitudes or dispositions to choose in certain ways in particular decision-making situations and more general dispositions. The analyst tries to identify, even quantify the conflicts, and then hopes that the client can resolve them in a manner more compatible with his or her "basic goals."

11.6 Some Bum Raps for Decision Analysis

A popular misconception is that decision analysis is unemotional, dehumanizing, and obsessive because it uses numbers and arithmetic in order to guide important life decisions. Isn't this turning over important human decisions "to a machine," sometimes literally a computer—which now picks our quarterbacks, our chief executive officers, and even our lovers? Aren't the "mathematicizers" of life, who admittedly have done well in the basic sciences, moving into a context where such uses of numbers are irrelevant and irreverent? Don't we suffer enough from the tyranny of numbers when our opportunities in life are controlled by numerical scores on aptitude tests and numbers entered on rating forms by interviewers and supervisors? In short, isn't the human spirit better expressed by intuitive choices than by analytic number crunching?

Our answer to all these concerns is an unqualified "no." There is absolutely nothing in the von Neumann and Morgenstern theory—or in this book—that requires the adoption of "inhumanly" stable or easily accessed values. In fact, the whole idea of utility is that it provides a measure of what is truly *personally* important to individuals reaching decisions. As presented here, the aim of analyzing expected utility is to help us achieve what is really important to us. As James March (1978) points out, one goal in life may be to discover what our values are. That goal might require action that is playful, or even arbitrary. Does such action violate the dictates of either rationality or expected utility theory? No. Upon examination, an individual valuing such an approach will be found to have a utility associated with the existential experimentation that follows from it. All that the decision analyst does is help to make this value explicit so that the individual can understand it and incorporate it into action in a noncontradictory manner.

Nor is decision analysis an obsessive, equivocating activity. In fact, some conclusions will mandate action rather than thought. For example, as mentioned earlier, there is a great deal more in von Neumann and Morgenstern's classic *Theory of Games and Economic Behavior* (1947) than has been presented here. One particularly intriguing section of that book concerns optimal play in poker games. There are 2,598,960 possible poker hands and, because no 2 of these hands are tied, drawing a particular hand is equivalent to drawing some number between 1 and 2,598,960. Since a hand is won by the person with the highest number, the question is what constitutes good betting strategy. Von Neumann and Morgenstern considered a simplified form of poker in which only two people play. Each person must ante, one person bets, and the other has the opportunity to either match the bet or raise it, at which point the first person may respond by matching the raise. What von Neumann and Morgenstern proved mathematically is that, according to the principle of maximizing expected utility, a player should either bet the maximum amount immediately or fold. (If the player is the first bettor, he or she may "check" or defer the bet until the other players act.) Our point is that this is a rigorous demonstration, within the context of the theory, that hesitant behavior is poor strategy. It implies the exact opposite of obsessing about a decision. In fact, the maximal strategy is to choose some number between 1 and 2,598,960 prior to looking at the value of the hand, bet the maximal amount if the value is above the chosen number, and otherwise not bet at all. In this context, dynamic decision making is supported. Absolutely nothing in the theory encourages people to obsess, procrastinate, or postpone.

References

Baron, J. (1988). *Thinking and deciding*. New York: Cambridge University Press.
Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes, 70,* 1–16.
Bentham, J. (1948). *An introduction to the principles of morals and legislations.* Oxford, UK: Blackwell. (Original work published 1789)
Clemen, R. T. (1996). *Making hard decisions: An introduction to decision analysis* (2nd ed.). Pacific Grove, CA: Duxbury Press.
Dawes, R. M., & Smith, T. L. (1985). Attitude and opinion measurement. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 1, pp. 509–566). New York: Random House.
Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75,* 643–669.
Freedman, D., Pisani, R., Purves, R., & Adhikari, A. (1991). *Statistics* (2nd ed.). New York: Norton.
Fryback, D. G., & Thornbury, J. R. (1976). Evaluation of a computerized Bayesian model for diagnosis of renal cysts versus tumor versus normal variant from exploratory urogram information. *Investigative Radiology, 11,* 102–111.
Hammond, J. S., Keeney, R. L., & Raiffa, H. (1999). *Smart choices: A practical guide to making better decisions.* Cambridge, MA: Harvard Business School Press.
Hastie, R., & Stasser, G. (2000). Computer simulation methods in social psychology. In C. M. Judd & H. Reis (Eds.), *Handbook of research methods in social psychology* (pp. 85–114). New York: Cambridge University Press.
Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass. *Reports on the negotiations of the Royal Saxon Society of Sciences, Mathematics and Physics Class, 53,* 1–64.
Kadane, J. B., Schervish, M. J., & Seidenfeld, T. (1999). *Rethinking the foundations of statistics.* Cambridge, UK: Cambridge University Press.
Lopes, L. L. (1994). Psychology and economics: Perspectives on risk, cooperation, and the marketplace. *Annual Review of Psychology, 45,* 197–227.
Lopes, L. L., & Oden, G. C. (1999). The role of aspiration level in risky choice: A comparison cumulative prospect theory and SP/A theory. *Journal of Mathematical Psychology, 43,* 286–313.
MacCrimmon, K. R., & Wehrung, D. A. (1986). *Taking risks: The management of uncertainty.* New York: Free Press.
March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics, 9,* 587–608.
Neumann, J. von, & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton, NJ: Princeton University Press.
Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

12

A Descriptive Decision Theory

In theory there is no difference between theory and practice. But in practice, there is.

—Jan L. A. van de Snepscheut and others

12.1 Non-expected Utility Theories

Both the economists' paradoxes and psychologists' experiments have repeatedly shown that subjective expected utility theory is not a valid descriptive theory of human behavior. Most efforts to create a more adequate descriptive theory have the basic form of the rational expectations principle (Section 2.3). Thus, they are sometimes called *non*-expected utility theories as a reminder that they are derived from the expected utility framework.

But why should we work within the general expected utility framework? First, the framework includes the ingredients that our intuitions and experience tell us are essential to deliberate decision making. Second, the framework provides a roughly accurate descriptive account of decision behaviors in many situations; some economists call it a *positive theory* because it relates inputs and outputs (psychologists might say stimuli and responses) in decision behavior to one another approximately correctly. And third, the framework captures the essence of rationality (as best our culture can define it), and it is likely that we are adapted to be approximately rational in our behaviors; our optimistic hypothesis is that people are at least half-smart in achieving their personal goals.