APPENDIX A: JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES*

Amos Tversky and Daniel Kahneman

Many decisions are based on beliefs concerning the likelihood of uncertain events such as the outcome of an election, the guilt of a defendant, or the future value of the dollar. These beliefs are usually expressed in statements such as "I think that \ldots ," "chances are \ldots ," "it is unlikely that \ldots ," and so forth. Occasionally, beliefs concerning uncertain events are expressed in numerical form as odds or subjective probabilities. What determines such beliefs? How do people assess the probability of an uncertain event or the value of an uncertain quantity? This article shows that people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors.

The subjective assessment of probability resembles the subjective assessment of physical quantities such as distance or size. These judgments are all based on data of limited validity, which are processed according to heuristic rules. For example, the apparent distance of an object is determined in part by its clarity. The more sharply the object is seen, the closer it appears to be. This rule has some validity, because in any given scene the more distant objects are seen less sharply than nearer objects. However, the reliance on this rule leads to systematic errors in the estimation of distance. Specifically, distances are often overestimated when visibility is poor because the contours of objects are blurred. On the other hand, distances are often underestimated when visibility is good because the objects are seen sharply. Thus, the reliance on clarity as an indication of distance leads to common biases. Such biases are also found in the intuitive judgment of probability. This article describes three heuristics that are employed to assess probabilities and to predict values. Biases to which these heuristics

*This article originally appeared in *Science*, vol. 185, 1974. The research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under contract N00014-73-C-0438 to the Oregon Research Institute, Eugene. Additional support for this research was provided by the Research and Development Authority of the Hebrew University, Jerusalem, Israel.

1

lead are enumerated, and the applied and theoretical implications of these observations are discussed.

REPRESENTATIVENESS

Many of the probabilistic questions with which people are concerned belong to one of the following types: What is the probability that object A belongs to class B? What is the probability that event A originates from process B? What is the probability that process B will generate event A? In answering such questions, people typically rely on the representativeness heuristic, in which probabilities are evaluated by the degree to which A is representative of B, that is, by the degree to which A resembles B. For example, when A is highly representative of B, the probability that A originates from B is judged to be high. On the other hand, if A is not similar to B, the probability that A originates from B is judged to be low.

For an illustration of judgment by representativeness, consider an individual who has been described by a former neighbor as follows: "Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail." How do people assess the probability that Steve is engaged in a particular occupation from a list of possibilities (for example, farmer, salesman, airline pilot, librarian, or physician)? How do people order these occupations from most to least likely? In the representativeness heuristic, the probability that Steve is a librarian, for example, is assessed by the degree to which he is representative of, or similar to, the stereotype of a librarian. Indeed, research with problems of this type has shown that people order the occupations by probability and by similarity in exactly the same way.¹ This approach to the judgment of probability leads to serious errors, because similarity, or representativeness, is not influenced by several factors that should affect judgments of probability.

Insensitivity to prior probability of outcomes. One of the factors that have no effect on representativeness but should have a major effect on probability is the prior probability, or baserate frequency, of the outcomes. In the case of Steve, for example, the fact that there are many more farmers than librarians in the population should enter into any reasonable estimate of the probability that Steve is a librarian rather than a farmer. Considerations of base-rate frequency, however, do not affect the similarity of Steve to the stereotypes of librarians and farmers. If people evaluate probability by representativeness, therefore, prior probabilities will be neglected. This hypothesis was tested in an experiment where prior probabilities were manipulated.² Subjects were shown brief personality descriptions of several individuals, allegedly sampled at random from a group of 100 professionals-engineers and lawyers. The subjects were asked to assess, for each description, the probability that it belonged to an engineer rather than to a lawyer. In one experimental condition, subjects were told that the group from which the descriptions had been drawn consisted of 70 engineers and 30 lawyers. In another condition, subjects were told that the group consisted of 30 engineers and 70 lawyers. The odds that any particular description belongs to an engineer rather than to a lawyer should be higher in the first condition, where there is a majority of engineers, than in the second condition, where there is a majority of lawyers. Specifically, it can be shown by applying Bayes' rule that the ratio of these odds should be $(.7/.3)^2$, or 5.44, for each description. In a sharp violation of Bayes' rule, the subjects in the two conditions produced essentially the same probability judgments. Apparently, subjects evaluated the likelihood that a particular description belonged to an engineer rather than to a lawyer by the degree to which this description was

representative of the two stereotypes, with little or no regard for the prior probabilities of the categories.

The subjects used prior probabilities correctly when they had no other information. In the absence of a personality sketch, they judged the probability that an unknown individual is an engineer to be .7 and .3, respectively, in the two base-rate conditions. However, prior probabilities were effectively ignored when a description was introduced, even when this description was totally uninformative. The responses to the following description illustrate this phenomenon:

Dick is a 30-year-old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues.

This description was intended to convey no information relevant to the question of whether Dick is an engineer or a lawyer. Consequently, the probability that Dick is an engineer should equal the proportion of engineers in the group, as if no description had been given. The subjects, however, judged the probability of Dick being an engineer to be .5 regardless of whether the stated proportion of engineers in the group was .7 or .3. Evidently, people respond differently when given no evidence and when given worthless evidence. When no specific evidence is given, prior probabilities are properly utilized; when worthless evidence is given, prior probabilities are ignored.³

Insensitivity to sample size. To evaluate the probability of obtaining a particular result in a sample drawn from a specified population, people typically apply the representativeness heuristic. That is, they assess the likelihood of a sample result, for example, that the average height in a random sample of ten men will be 6 feet, by the similarity of this result to the corresponding parameter (that is, to the average height in the population of men). The similarity of a sample statistic to a population parameter does not depend on the size of the sample. Consequently, if probabilities are assessed by representativeness, then the judged probability of a sample statistic will be essentially independent of sample size. Indeed, when subjects assessed the distributions of average height for samples of various sizes, they produced identical distributions. For example, the probability of obtaining an average height greater than 6 feet was assigned the same value for samples of 1,000, 100, and 10 men.⁴ Moreover, subjects failed to appreciate the role of sample size even when it was emphasized in the formulation of the problem. Consider the following question:

A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. However, the exact percentage varies from day to day. Sometimes it may be higher than 50%, sometimes lower.

For a period of 1 year, each hospital recorded the days on which more than 60% of the babies born were boys. Which hospital do you think recorded more such days?

The larger hospital (21)

The smaller hospital (21)

About the same (that is, within 5% of each other) (53)

The values in parentheses are the number of undergraduate students who chose each answer. Most subjects judged the probability of obtaining more than 60% boys to be the same in the small and in the large hospital, presumably because these events are described by the same sta-

3

tistic and are therefore equally representative of the general population. In contrast, sampling theory entails that the expected number of days on which more than 60% of the babies are boys is much greater in the small hospital than in the large one, because a large sample is less likely to stray from 50%. This fundamental notion of statistics is evidently not part of people's repertoire of intuitions.

A similar insensitivity to sample size has been reported in judgments of posterior probability, that is, of the probability that a sample has been drawn from one population rather than from another. Consider the following example:

Imagine an urn filled with balls, of which $\frac{2}{3}$ are of one color and $\frac{1}{3}$ of another. One individual has drawn 5 balls from the urn, and found that 4 were red and 1 was white. Another individual has drawn 20 balls and found that 12 were red and 8 were white. Which of the two individuals should feel more confident that the urn contains $\frac{2}{3}$ red balls and $\frac{1}{3}$ white balls, rather than the opposite? What odds should each individual give?

In this problem, the correct posterior odds are 8 to 1 for the 4:1 sample and 16 to 1 for the 12:8 sample, assuming equal prior probabilities. However, most people feel that the first sample provides much stronger evidence for the hypothesis that the urn is predominantly red, because the proportion of red balls is larger in the first than in the second sample. Here again, intuitive judgments are dominated by the sample proportion and are essentially unaffected by the size of the sample, which plays a crucial role in the determination of the actual posterior odds.⁵ In addition, intuitive estimates of posterior odds are far less extreme than the correct values. The underestimation of the impact of evidence has been observed repeatedly in problems of this type.⁶ It has been labeled "conservatism."

Misconceptions of chance. People expect that a sequence of events generated by a random process will represent the essential characteristics of that process even when the sequence is short. In considering tosses of a coin for heads or tails, for example, people regard the sequence H-T-H-T-T-H to be more likely than the sequence H-H-H-T-T-T, which does not appear random, and also more likely than the sequence H-H-H-H-T-H, which does not represent the fairness of the coin.⁷ Thus, people expect that the essential characteristics of the process will be represented, not only globally in the entire sequence, but also locally in each of its parts. A locally representative sequence, however, deviates systematically from chance expectation: it contains too many alternations and too few runs. Another consequence of the belief in local representativeness is the well-known gambler's fallacy. After observing a long run of red on the roulette wheel, for example, most people erroneously believe that black is now due, presumably because the occurrence of black will result in a more representative sequence than the occurrence of an additional red. Chance is commonly viewed as a self-correcting process in which a deviation in one direction induces a deviation in the opposite direction to restore the equilibrium. In fact, deviations are not "corrected" as a chance process unfolds, they are merely diluted.

Misconceptions of chance are not limited to naive subjects. A study of the statistical intuitions of experienced research psychologists⁸ revealed a lingering belief in what may be called the "law of small numbers," according to which even small samples are highly representative of the populations from which they are drawn. The responses of these investigators reflected the expectation that a valid hypothesis about a population will be represented by a statistically significant result in a sample with little regard for its size. As a consequence, the researchers put too much faith in the results of small samples and grossly overestimated the replicability of such

results. In the actual conduct of research, this bias leads to the selection of samples of inadequate size and to overinterpretation of findings.

Insensitivity to predictability. People are sometimes called upon to make such numerical predictions as the future value of a stock, the demand for a commodity, or the outcome of a football game. Such predictions are often made by representativeness. For example, suppose one is given a description of a company and is asked to predict its future profit. If the description of the company is very favorable, a very high profit will appear most representative of that description; if the description is mediocre, a mediocre performance will appear most representative. The degree to which the description is favorable is unaffected by the reliability of that description or by the degree to which it permits accurate prediction. Hence, if people predict solely in terms of the favorableness of the description, their predictions will be insensitive to the reliability of the evidence and to the expected accuracy of the prediction.

This mode of judgment violates the normative statistical theory in which the extremeness and the range of predictions are controlled by considerations of predictability. When predictability is nil, the same prediction should be made in all cases. For example, if the descriptions of companies provide no information relevant to profit, then the same value (such as average profit) should be predicted for all companies. If predictability is perfect, of course, the values predicted will match the actual values and the range of predictions will equal the range of outcomes. In general, the higher the predictability, the wider the range of predicted values.

Several studies of numerical prediction have demonstrated that intuitive predictions violate this rule, and that subjects show little or no regard for considerations of predictability.⁹ In one of these studies, subjects were presented with several paragraphs, each describing the performance of a student teacher during a particular practice lesson. Some subjects were asked to evaluate the quality of the lesson described in the paragraph in percentile scores, relative to a specified population. Other subjects were asked to predict, also in percentile scores, the standing of each student teacher 5 years after the practice lesson. The judgments made under the two conditions were identical. That is, the prediction of a remote criterion (success of a teacher after 5 years) was identical to the evaluation of the information on which the prediction was based (the quality of the practice lesson). The students who made these predictions were undoubtedly aware of the limited predictability of teaching competence on the basis of a single trial lesson 5 years earlier; nevertheless, their predictions were as extreme as their evaluations.

The illusion of validity. As we have seen, people often predict by selecting the outcome (for example, an occupation) that is most representative of the input (for example, the description of a person). The confidence they have in their prediction depends primarily on the degree of representativeness (that is, on the quality of the match between the selected outcome and the input) with little or no regard for the factors that limit predictive accuracy. Thus, people express great confidence in the prediction that a person is a librarian when given a description of his personality which matches the stereotype of librarians, even if the description is scanty, unreliable, or outdated. The unwarranted confidence which is produced by a good fit between the predicted outcome and the input information may be called the illusion of validity. This illusion persists even when the judge is aware of the factors that limit the accuracy of his predictions. It is a common observation that psychologists who conduct selection interviews often experience considerable confidence in their predictions, even when they know of the vast literature that shows selection interviews to be highly fallible. The continued reliance on the clinical interview for selection, despite repeated demonstrations of its inadequacy, amply attests to the strength of this effect.

The internal consistency of a pattern of inputs is a major determinant of one's confidence

in predictions based on these inputs. For example, people express more confidence in predicting the final grade point average of a student whose first-year record consists entirely of B's than in predicting the grade point average of a student whose first-year record includes many A's and C's. Highly consistent patterns are most often observed when the input variables are highly redundant or correlated. Hence, people tend to have great confidence in predictions based on redundant input variables. However, an elementary result in the statistics of correlation asserts that, given input variables of stated validity, a prediction based on several such inputs can achieve higher accuracy when they are independent of each other than when they are redundant or correlated. Thus, redundancy among inputs decreases accuracy even as it increases confidence, and people are often confident in predictions that are quite likely to be off the mark.¹⁰

Misconceptions of regression. Suppose a large group of children has been examined on two equivalent versions of an aptitude test. If one selects ten children from among those who did best on one of the two versions, he will usually find their performance on the second version to be somewhat disappointing. Conversely, if one selects ten children from among those who did worst on one version, they will be found, on the average, to do somewhat better on the other version. More generally, consider two variables X and Y which have the same distribution. If one selects individuals whose average X score deviates from the mean of X by k units, then the average of their Y scores will usually deviate from the mean of Y by less than k units. These observations illustrate a general phenomenon known as regression toward the mean, which was first documented by Galton more than 100 years ago.

In the normal course of life, one encounters many instances of regression toward the mean, in the comparison of the height of fathers and sons, of the intelligence of husbands and wives, or of the performance of individuals on consecutive examinations. Nevertheless, people do not develop correct intuitions about this phenomenon. First, they do not expect regression in many contexts where it is bound to occur. Second, when they recognize the occurrence of regression, they often invent spurious causal explanations for it.¹¹ We suggest that the phenomenon of regression remains elusive because it is incompatible with the belief that the predicted outcome should be maximally representative of the input, and, hence, that the value of the outcome variable should be as extreme as the value of the input variable.

The failure to recognize the import of regression can have pernicious consequences, as illustrated by the following observation.¹² In a discussion of flight training, experienced instructors noted that praise for an exceptionally smooth landing is typically followed by a poorer landing on the next try, while harsh criticism after a rough landing is usually followed by an improvement on the next try. The instructors concluded that verbal rewards are detrimental to learning, while verbal punishments are beneficial, contrary to accepted psychological doctrine. This conclusion is unwarranted because of the presence of regression toward the mean. As in other cases of repeated examination, an improvement will usually follow a poor performance and a deterioration will usually follow an outstanding performance, even if the instructor does not respond to the trainee's achievement on the first attempt. Because the instructors had praised their trainees after good landings and admonished them after poor ones, they reached the erroneous and potentially harmful conclusion that punishment is more effective than reward.

Thus, the failure to understand the effect of regression leads one to overestimate the effectiveness of punishment and to underestimate the effectiveness of reward. In social interaction, as well as in training, rewards are typically administered when performance is good, and punishments are typically administered when performance is poor. By regression alone, therefore, behavior is most likely to improve after punishment and most likely to deteriorate after reward.

Consequently, the human condition is such that, by chance alone, one is most often rewarded for punishing others and most often punished for rewarding them. People are generally not aware of this contingency. In fact, the elusive role of regression in determining the apparent consequences of reward and punishment seems to have escaped the notice of students of this area.

AVAILABILITY

There are situations in which people assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind. For example, one may assess the risk of heart attack among middle-aged people by recalling such occurrences among one's acquaintances. Similarly, one may evaluate the probability that a given business venture will fail by imagining various difficulties it could encounter. This judgmental heuristic is called availability. Availability is a useful clue for assessing frequency or probability, because instances of large classes are usually recalled better and faster than instances of less frequent classes. However, availability is affected by factors other than frequency and probability. Consequently, the reliance on availability leads to predictable biases, some of which are illustrated below.

Biases due to the retrievability of instances. When the size of a class is judged by the availability of its instances, a class whose instances are easily retrieved will appear more numerous than a class of equal frequency whose instances are less retrievable. In an elementary demonstration of this effect, subjects heard a list of well-known personalities of both sexes and were subsequently asked to judge whether the list contained more names of men than of women. Different lists were presented to different groups of subjects. In some of the lists the men were relatively more famous than the women, and in others the women were relatively more famous than the men. In each of the lists, the subjects erroneously judged that the class (sex) that had the more famous personalities was the more numerous.¹³

In addition to familiarity, there are other factors, such as salience, which affect the retrievability of instances. For example, the impact of seeing a house burning on the subjective probability of such accidents is probably greater than the impact of reading about a fire in the local paper. Furthermore, recent occurrences are likely to be relatively more available than earlier occurrences. It is a common experience that the subjective probability of traffic accidents rises temporarily when one sees a car overturned by the side of the road.

Biases due to the effectiveness of a search set. Suppose one samples a word (of three letters or more) at random from an English text. Is it more likely that the word starts with r or that r is the third letter? People approach this problem by recalling words that begin with r (*road*) and words that have r in the third position (*car*) and assess the relative frequency by the ease with which words of the two types come to mind. Because it is much easier to search for words by their first letter than by their third letter, most people judge words that begin with a given consonant to be more numerous than words in which the same consonant appears in the third position. They do so even for consonants, such as r or k, that are more frequent in the third position than in the first.¹⁴

Different tasks elicit different search sets. For example, suppose you are asked to rate the frequency with which abstract words (*thought*, *love*) and concrete words (*door*, *water*) appear in written English. A natural way to answer this question is to search for contexts in which the word could appear. It seems easier to think of contexts in which an abstract concept is mentioned (love in love stories) than to think of contexts in which a concrete word (such as *door*) is mentioned. If the frequency of words is judged by the availability of the contexts in which they

appear, abstract words will be judged as relatively more numerous than concrete words. This bias has been observed in a recent study¹⁵ which showed that the judged frequency of occurrence of abstract words was much higher than that of concrete words, equated in objective frequency. Abstract words were also judged to appear in a much greater variety of contexts than concrete words.

Biases of imaginability. Sometimes one has to assess the frequency of a class whose instances are not stored in memory but can be generated according to a given rule. In such situations, one typically generates several instances and evaluates frequency or probability by the ease with which the relevant instances can be constructed. However, the ease of constructing instances does not always reflect their actual frequency, and this mode of evaluation is prone to biases. To illustrate, consider a group of 10 people who form committees of *k* members, $2 \le k \le 8$. How many different committees of *k* members can be formed? The correct answer to this problem is given by the binomial coefficient (10/*k*) which reaches a maximum of 252 for k = 5. Clearly, the number of committees of *k* members equals the number of committees of (10 - *k*) members, because any committee of *k* members defines a unique group of (10 - *k*) nonmembers.

One way to answer this question without computation is to mentally construct committees of k members and to evaluate their number by the ease with which they come to mind. Committees of few members, say 2, are more available than committees of many members, say 8. The simplest scheme for the construction of committees is a partition of the group into disjoint sets. One readily sees that it is easy to construct five disjoint committees of 2 members, while it is impossible to generate even two disjoint committees of 8 members. Consequently, if frequency is assessed by imaginability, or by availability for construction, the small committees will appear more numerous than larger committees, in contrast to the correct bell-shaped function. Indeed, when naive subjects were asked to estimate the number of distinct committees of various sizes, their estimates were a decreasing monotonic function of committee size.¹⁶ For example, the median estimate of the number of committees of 2 members was 20 (the correct answer is 45 in both cases).

Imaginability plays an important role in the evaluation of probabilities in real-life situations. The risk involved in an adventurous expedition, for example, is evaluated by imagining contingencies with which the expedition is not equipped to cope. If many such difficulties are vividly portrayed, the expedition can be made to appear exceedingly dangerous, although the ease with which disasters are imagined need not reflect their actual likelihood. Conversely, the risk involved in an undertaking may be grossly underestimated if some possible dangers are either difficult to conceive of, or simply do not come to mind.

Illusory correlation. Chapman and Chapman¹⁷ have described an interesting bias in the judgment of the frequency with which two events co-occur. They presented naive judges with information concerning several hypothetical mental patients. The data for each patient consisted of a clinical diagnosis and a drawing of a person made by the patient. Later the judges estimated the frequency with which each diagnosis (such as paranoia or suspiciousness) had been accompanied by various features of the drawing (such as peculiar eyes). The subjects markedly overestimated the frequency of co-occurrence of natural associates, such as suspiciousness and peculiar eyes. This effect was labeled illusory correlation. In their erroneous judgments of the data to which they had been exposed, naive subjects "rediscovered" much of the common, but unfounded, clinical lore concerning the interpretation of the draw-a-person test. The illusory correlation effect was extremely resistant to contradictory data. It persisted

even when the correlation between symptom and diagnosis was actually negative, and it prevented the judges from detecting relationships that were in fact present.

Availability provides a natural account for the illusory-correlation effect. The judgment of how frequently two events co-occur could be based on the strength of the associative bond between them. When the association is strong, one is likely to conclude that the events have been frequently paired. Consequently, strong associates will be judged to have occurred together frequently. According to this view, the illusory correlation between suspiciousness and peculiar drawing of the eyes, for example, is due to the fact that suspiciousness is more readily associated with the eyes than with any other part of the body.

Lifelong experience has taught us that, in general, instances of large classes are recalled better and faster than instances of less frequent classes; that likely occurrences are easier to imagine than unlikely ones; and that the associative connections between events are strengthened when the events frequently co-occur. As a result, man has at his disposal a procedure (the availability heuristic) for estimating the numerosity of a class, the likelihood of an event, or the frequency of co-occurrences, by the ease with which the relevant mental operations of retrieval, construction, or association can be performed. However, as the preceding examples have demonstrated, this valuable estimation procedure results in systematic errors.

ADJUSTMENT AND ANCHORING

In many situations, people make estimates by starting from an initial value that is adjusted to yield the final answer. The initial value, or starting point, may be suggested by the formulation of the problem, or it may be the result of a partial computation. In either case, adjustments are typically insufficient.¹⁸ That is, different starting points yield different estimates, which are biased toward the initial values. We call this phenomenon anchoring.

Insufficient adjustment. In a demonstration of the anchoring effect, subjects were asked to estimate various quantities, stated in percentages (for example, the percentage of African countries in the United Nations). For each quantity, a number between 0 and 100 was determined by spinning a wheel of fortune in the subjects' presence. The subjects were instructed to indicate first whether that number was higher or lower than the value of the quantity, and then to estimate the value of the quantity by moving upward or downward from the given number. Different groups were given different numbers for each quantity, and these arbitrary numbers had a marked effect on estimates. For example, the median estimates of the percentage of African countries in the United Nations were 25 and 45 for groups that received 10 and 65, respectively, as starting points. Payoffs for accuracy did not reduce the anchoring effect.

Anchoring occurs not only when the starting point is given to the subject, but also when the subject bases his estimate on the result of some incomplete computation. A study of intuitive numerical estimation illustrates this effect. Two groups of high school students estimated, within 5 seconds, a numerical expression that was written on the blackboard. One group estimated the product

 $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

while another group estimated the product

$$1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$$

105838_01_1-484_r5sk.indd

To rapidly answer such questions, people may perform a few steps of computation and estimate the product by extrapolation or adjustment. Because adjustments are typically insufficient, this procedure should lead to underestimation. Furthermore, because the result of the first few steps of multiplication (performed from left to right) is higher in the descending sequence than in the ascending sequence, the former expression should be judged larger than the latter. Both predictions were confirmed. The median estimate for the ascending sequence was 512, while the median estimate for the descending sequence was 2,250. The correct answer is 40,320.

Biases in the evaluation of conjunctive and disjunctive events. In a recent study by Bar-Hillel¹⁹ subjects were given the opportunity to bet on one of two events. Three types of events were used: (i) simple events, such as drawing a red marble from a bag containing 50% red marbles and 50% white marbles; (ii) conjunctive events, such as drawing a red marble seven times in succession, with replacement, from a bag containing 90% red marbles and 10% white marbles; and (iii) disjunctive events, such as drawing a red marble at least once in seven successive tries, with replacement, from a bag containing 10% red marbles and 90% white marbles. In this problem, a significant majority of subjects preferred to bet on the conjunctive event (the probability of which is .48) rather than on the simple event (the probability of which is .50). Subjects also preferred to bet on the simple event rather than on the disjunctive event, which has a probability of .52. Thus, most subjects bet on the less likely event in both comparisons. This pattern of choices illustrates a general finding. Studies of choice among gambles and of judgments of probability indicate that people tend to overestimate the probability of conjunctive events²⁰ and to underestimate the probability of disjunctive events. These biases are readily explained as effects of anchoring. The stated probability of the elementary event (success at any one stage) provides a natural starting point for the estimation of the probabilities of both conjunctive and disjunctive events. Since adjustment from the starting point is typically insufficient, the final estimates remain too close to the probabilities of the elementary events in both cases. Note that the overall probability of a conjunctive event is lower than the probability of each elementary event, whereas the overall probability of a disjunctive event is higher than the probability of each elementary event. As a consequence of anchoring, the overall probability will be overestimated in conjunctive problems and underestimated in disjunctive problems.

Biases in the evaluation of compound events are particularly significant in the context of planning. The successful completion of an undertaking, such as the development of a new product, typically has a conjunctive character: for the undertaking to succeed, each of a series of events must occur. Even when each of these events is very likely, the overall probability of success can be quite low if the number of events is large. The general tendency to overestimate the probability of conjunctive events leads to unwarranted optimism in the evaluation of the likelihood that a plan will succeed or that a project will be completed on time. Conversely, disjunctive structures are typically encountered in the evaluation of risks. A complex system, such as a nuclear reactor or a human body, will malfunction if any of its essential components fails. Even when the likelihood of failure in each component is slight, the probability of an overall failure can be high if many components are involved. Because of anchoring, people will tend to underestimate the probabilities of failure in complex systems. Thus, the direction of the an-choring bias can sometimes be inferred from the structure of the event. The chain-like structure of conjunctions leads to overestimation, the funnel-like structure of disjunctions leads to underestimation.

Anchoring in the assessment of subjective probability distributions. In decision analysis, ex-

perts are often required to express their beliefs about a quantity, such as the value of the Dow Jones average on a particular day, in the form of a probability distribution. Such a distribution is usually constructed by asking the person to select values of the quantity that correspond to specified percentiles of his subjective probability distribution. For example, the judge may be asked to select a number, X_{90} , such that his subjective probability that this number will be higher than the value of the Dow Jones average is .90. That is, he should select the value X_{90} so that he is just willing to accept 9 to 1 odds that the Dow Jones average will not exceed it. A subjective probability distribution for the value of the Dow Jones average can be constructed from several such judgments corresponding to different percentiles.

By collecting subjective probability distributions for many different quantities, it is possible to test the judge for proper calibration. A judge is properly (or externally) calibrated in a set of problems if exactly Π % of the true values of the assessed quantities falls below his stated values of X_{Π} . For example, the true values should fall below X_{01} for 1% of the quantities and above X_{99} for 1% of the quantities. Thus, the true values should fall in the confidence interval between X_{01} and X_{99} on 98% of the problems.

Several investigators²¹ have obtained probability distributions for many quantities from a large number of judges. These distributions indicated large and systematic departures from proper calibration. In most studies, the actual values of the assessed quantities are either smaller than X_{0l} or greater than X_{99} for about 30% of the problems. That is, the subjects state overly narrow confidence intervals which reflect more certainty than is justified by their knowledge about the assessed quantities. This bias is common to naive and to sophisticated subjects, and it is not eliminated by introducing proper scoring rules, which provide incentives for external calibration. This effect is attributable, in part at least, to anchoring.

To select X_{90} for the value of the Dow Jones average, for example, it is natural to begin by thinking about one's best estimate of the Dow Jones and to adjust this value upward. If this adjustment—like most others—is insufficient, then X_{90} will not be sufficiently extreme. A similar anchoring effect will occur in the selection of X_{10} , which is presumably obtained by adjusting one's best estimate downward. Consequently, the confidence interval between X_{10} and X_{90} will be too narrow, and the assessed probability distribution will be too tight. In support of this interpretation it can be shown that subjective probabilities are systematically altered by a procedure in which one's best estimate does not serve as an anchor.

Subjective probability distributions for a given quantity (the Dow Jones average) can be obtained in two different ways: (i) by asking the subject to select values of the Dow Jones that correspond to specified percentiles of his probability distribution and (ii) by asking the subject to assess the probabilities that the true value of the Dow Jones will exceed some specified values. The two procedures are formally equivalent and should yield identical distributions. However, they suggest different modes of adjustment from different anchors. In procedure (i), the natural starting point is one's best estimate of the quantity. In procedure (ii), on the other hand, the subject may be anchored on the value stated in the question. Alternatively, he may be anchored on even odds, or a 50–50 chance, which is a natural starting point in the estimation of likelihood. In either case, procedure (ii) should yield less extreme odds than procedure (i).

To contrast the two procedures, a set of 24 quantities (such as the air distance from New Delhi to Peking) was presented to a group of subjects who assessed either X_{10} or X_{90} for each problem. Another group of subjects received the median judgment of the first group for each of the 24 quantities. They were asked to assess the odds that each of the given values exceeded the

()

APPENDIX A

true value of the relevant quantity. In the absence of any bias, the second group should retrieve the odds specified to the first group, that is, 9:1. However, if even odds or the stated value serve as anchors, the odds of the second group should be less extreme, that is, closer to 1:1. Indeed, the median odds stated by this group, across all problems, were 3:1. When the judgments of the two groups were tested for external calibration, it was found that subjects in the first group were too extreme, in accord with earlier studies. The events that they defined as having a probability of .10 actually obtained in 24% of the cases. In contrast, subjects in the second group were too conservative. Events to which they assigned an average probability of .34 actually obtained in 26% of the cases. These results illustrate the manner in which the degree of calibration depends on the procedure of elicitation.

DISCUSSION

This article has been concerned with cognitive biases that stem from the reliance on judgmental heuristics. These biases are not attributable to motivational effects such as wishful thinking or the distortion of judgments by payoffs and penalties. Indeed, several of the severe errors of judgment reported earlier occurred despite the fact that subjects were encouraged to be accurate and were rewarded for the correct answers.²²

The reliance on heuristics and the prevalence of biases are not restricted to laymen. Experienced researchers are also prone to the same biases—when they think intuitively. For example, the tendency to predict the outcome that best represents the data, with insufficient regard for prior probability, has been observed in the intuitive judgments of individuals who have had extensive training in statistics.²³ Although the statistically sophisticated avoid elementary errors, such as the gambler's fallacy, their intuitive judgments are liable to similar fallacies in more intricate and less transparent problems.

It is not surprising that useful heuristics such as representativeness and availability are retained, even though they occasionally lead to errors in prediction or estimation. What is perhaps surprising is the failure of people to infer from lifelong experience such fundamental statistical rules as regression toward the mean, or the effect of sample size on sampling variability. Although everyone is exposed, in the normal course of life, to numerous examples from which these rules could have been induced, very few people discover the principles of sampling and regression on their own. Statistical principles are not learned from everyday experience because the relevant instances are not coded appropriately. For example, people do not discover that successive lines in a text differ more in average word length than do successive pages, because they simply do not attend to the average word length of individual lines or pages. Thus, people do not learn the relation between sample size and sampling variability, although the data for such learning are abundant.

The lack of an appropriate code also explains why people usually do not detect the biases in their judgments of probability. A person could conceivably learn whether his judgments are externally calibrated by keeping a tally of the proportion of events that actually occur among those to which he assigns the same probability. However, it is not natural to group events by their judged probability. In the absence of such grouping it is impossible for an individual to discover, for example, that only 50% of the predictions to which he has assigned a probability of .9 or higher actually came true.

The empirical analysis of cognitive biases has implications for the theoretical and applied role of judged probabilities. Modern decision theory²⁴ regards subjective probability as the

quantified opinion of an idealized person. Specifically, the subjective probability of a given event is defined by the set of bets about this event that such a person is willing to accept. An internally consistent, or coherent, subjective probability measure can be derived for an individual if his choices among bets satisfy certain principles, that is, the axioms of the theory. The derived probability is subjective in the sense that different individuals are allowed to have different probabilities for the same event. The major contribution of this approach is that it provides a rigorous subjective interpretation of probability that is applicable to unique events and is embedded in a general theory of rational decision.

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.²⁵

The inherently subjective nature of probability has led many students to the belief that coherence, or internal consistency, is the only valid criterion by which judged probabilities should be evaluated. From the standpoint of the formal theory of subjective probability, any set of internally consistent probability judgments is as good as any other. This criterion is not entirely satisfactory, because an internally consistent set of subjective probabilities can be incompatible with other beliefs held by the individual. Consider a person whose subjective probabilities for all possible outcomes of a coin-tossing game reflect the gambler's fallacy. That is, his estimate of the probability of tails on a particular toss increases with the number of consecutive heads that preceded that toss. The judgments of such a person could be internally consistent and therefore acceptable as adequate subjective probabilities according to the criterion of the formal theory. These probabilities, however, are incompatible with the generally held belief that a coin has no memory and is therefore incapable of generating sequential dependencies. For judged probabilities to be considered adequate, or rational, internal consistency is not enough. The judgments must be compatible with the entire web of beliefs held by the individual. Unfortunately, there can be no simple formal procedure for assessing the compatibility of a set of probability judgments with the judge's total system of beliefs. The rational judge will nevertheless strive for compatibility, even though internal consistency is more easily achieved and assessed. In particular, he will attempt to make his probability judgments compatible with his knowledge about the subject matter, the laws of probability, and his own judgmental heuristics and biases.

SUMMARY

This article described three heuristics that are employed in making judgments under uncertainty: (i) representativeness, which is usually employed when people are asked to judge the probability that an object or event A belongs to class or process B; (ii) availability of instances or scenarios, which is often employed when people are asked to assess the frequency of a class or the plausibility of a particular development; and (iii) adjustment from an anchor, which is usually employed in numerical prediction when a relevant value is available. These heuristics are highly economical and usually effective, but they lead to systematic and predictable errors. A better understanding of these heuristics and of the biases to which they lead could improve judgments and decisions in situations of uncertainty.

NOTES

- D. Kahneman and A. Tversky, "On the Psychology of Prediction," *Psychological Review* 80 (1973): 237–51.
- 2. Ibid.
- 3. Ibid.
- D. Kahneman and A. Tversky, "Subjective Probability: A Judgment of Representativeness," Cognitive Psychology 3 (1972): 430–54.
- 5. Ibid.
- 6. W. Edwards, "Conservatism in Human Information Processing," in *Formal Representation* of *Human Judgment*, ed. B. Kleinmuntz (New York: Wiley, 1968), 17–52.
- 7. Kahneman and Tversky, "Subjective Probability."
- 8. A. Tversky and D. Kahneman, "Belief in the Law of Small Numbers," *Psychological Bulletin* 76 (1971): 105–10.
- 9. Kahneman and Tversky, "On the Psychology of Prediction."
- 10. Ibid.
- 11. Ibid.
- 12. Ibid.
- A. Tversky and D. Kahneman, "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology* 5 (1973): 207–32.
- 14. Ibid.
- R. C. Galbraith and B. J. Underwood, "Perceived Frequency of Concrete and Abstract Words," *Memory & Cognition* 1 (1973): 56–60.
- 16. Tversky and Kahneman, "Availability."
- L. J. Chapman and J. P. Chapman, "Genesis of Popular but Erroneous Psychodiagnostic Observations," *Journal of Abnormal Psychology* 73 (1967): 193–204; L. J. Chapman and J. P. Chapman, "Illusory Correlation as an Obstacle to the Use of Valid Psychodiagnostic Signs," *Journal of Abnormal Psychology* 74 (1969): 271–80.
- P. Slovic and S. Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," *Organizational Behavior & Human Performance* 6 (1971): 649–744.
- 19. M. Bar-Hillel, "On the Subjective Probability of Compound Events," Organizational Behavior & Human Performance 9 (1973): 396–406.
- J. Cohen, E. I. Chesnick, and D. Haran, "A Confirmation of the Inertial-Ψ Effect in Sequential Choice and Decision," *British Journal of Psychology* 63 (1972): 41–46.
- M. Alpert and H. Raiffa, unpublished manuscript; C. A. Stael von Holstein, "Two Techniques for Assessment of Subjective Probability Distributions: An Experimental Study," *Acta Psychologica* 35 (1971): 478–94; R. L. Winkler, "The Assessment of Prior Distributions in Bayesian Analysis," *Journal of the American Statistical Association* 62 (1967): 776–800.
- 22. Kahneman and Tversky, "Subjective Probability"; Tversky and Kahneman, "Availability."
- 23. Kahneman and Tversky, "On the Psychology of Prediction"; Tversky and Kahneman, "Belief in the Law of Small Numbers."
- 24. L. J. Savage, The Foundations of Statistics (New York: Wiley, 1954).
- Ibid.; B. de Finetti, "Probability: Interpretations," in *International Encyclopedia of the Social Sciences*, ed. D. E. Sills, vol. 12 (New York: Macmillan, 1968), 496–505.

APPENDIX B: CHOICES, VALUES, AND FRAMES*

Daniel Kahneman and Amos Tversky

ABSTRACT: We discuss the cognitive and the psychophysical determinants of choice in risky and riskless contexts. The psychophysics of value induce risk aversion in the domain of gains and risk seeking in the domain of losses. The psychophysics of chance induce overweighting of sure things and of improbable events, relative to events of moderate probability. Decision problems can be described or framed in multiple ways that give rise to different preferences, contrary to the invariance criterion of rational choice. The process of mental accounting, in which people organize the outcomes of transactions, explains some anomalies of consumer behavior. In particular, the acceptability of an option can depend on whether a negative outcome is evaluated as a cost or as an uncompensated loss. The relation between decision values and experience values is discussed.

Making decisions is like speaking prose—people do it all the time, knowingly or unknowingly. It is hardly surprising, then, that the topic of decision making is shared by many disciplines, from mathematics and statistics, through economics and political science, to sociology and psychology. The study of decisions addresses both normative and descriptive questions. The normative analysis is concerned with the nature of rationality and the logic of decision making. The descriptive analysis, in contrast, is concerned with people's beliefs and preferences as they are, not as they should be. The tension between normative and descriptive considerations characterizes much of the study of judgment and choice.

Analyses of decision making commonly distinguish risky and riskless choices. The paradigmatic example of decision under risk is the acceptability of a gamble that yields monetary outcomes with specified probabilities. A typical riskless decision concerns the acceptability of a transaction in which a good or a service is exchanged for money or labor. In the first part of this article we present an analysis of the cognitive and psychophysical factors that determine the value of risky prospects. In the second part we extend this analysis to transactions and trades.

*This article was originally presented as a Distinguished Scientific Contributions Award address at the American Psychological Association meeting, August 1983. This work was supported by grant NR 197-058 from the U.S. Office of Naval Research. Originally published in *American Psychologist*, vol. 34, 1984.

105838_01_1-484_r5sk.indd

RISKY CHOICE

Risky choices, such as whether or not to take an umbrella and whether or not to go to war, are made without advance knowledge of their consequences. Because the consequences of such actions depend on uncertain events such as the weather or the opponent's resolve, the choice of an act may be construed as the acceptance of a gamble that can yield various outcomes with different probabilities. It is therefore natural that the study of decision making under risk has focused on choices between simple gambles with monetary outcomes and specified probabilities, in the hope that these simple problems will reveal basic attitudes toward risk and value.

We shall sketch an approach to risky choice that derives many of its hypotheses from a psychophysical analysis of responses to money and to probability. The psychophysical approach to decision making can be traced to a remarkable essay that Daniel Bernoulli published in 1738 (Bernoulli 1954) in which he attempted to explain why people are generally averse to risk and why risk aversion decreases with increasing wealth. To illustrate risk aversion and Bernoulli's analysis, consider the choice between a prospect that offers an 85% chance to win \$1,000 (with a 15% chance to win nothing) and the alternative of receiving \$800 for sure. A large majority of people prefer the sure thing over the gamble, although the gamble has higher (mathematical) expectation. The expectation of a monetary gamble is a weighted average, where each possible outcome is weighted by its probability of occurrence. The expectation of the gamble in this example is .85 x \$1,000 + .15 x \$0 = \$850, which exceeds the expectation of \$800 associated with the sure thing. The preference for the sure gain is an instance of risk aversion. In general, a preference for a sure outcome over a gamble that has higher or equal expectation is called risk averse, and the rejection of a sure thing in favor of a gamble of lower or equal expectation is called risk seeking.

Bernoulli suggested that people do not evaluate prospects by the expectation of their monetary outcomes, but rather by the expectation of the subjective value of these outcomes. The subjective value of a gamble is again a weighted average, but now it is the subjective value of each outcome that is weighted by its probability. To explain risk aversion within this framework, Bernoulli proposed that subjective value, or utility, is a concave function of money. In such a function, the difference between the utilities of \$200 and \$100, for example, is greater than the utility difference between \$1,200 and \$1,100. It follows from concavity that the subjective value attached to a gain of \$800 is more than 80% of the value of a gain of \$1,000. Consequently, the concavity of the utility function entails a risk averse preference for a sure gain of \$800 over an 80% chance to win \$1,000, although the two prospects have the same monetary expectation.

It is customary in decision analysis to describe the outcomes of decisions in terms of total wealth. For example, an offer to bet \$20 on the toss of a fair coin is represented as a choice between an individual's current wealth W and an even chance to move to W + \$20 or to W - \$20. This representation appears psychologically unrealistic: People do not normally think of relatively small outcomes in terms of states of wealth but rather in terms of gains, losses, and neutral outcomes (such as the maintenance of the status quo). If the effective carriers of subjective value are changes of wealth rather than ultimate states of wealth, as we propose, the psychophysical analysis of outcomes should be applied to gains and losses rather than to total assets. This assumption plays a central role in a treatment of risky choice that we called prospect theory (Kahneman and Tversky 1979). Introspection as well as psychophysical measurements suggest

that subjective value is a concave function of the size of a gain. The same generalization applies to losses as well. The difference in subjective value between a loss of \$200 and a loss of \$100 appears greater than the difference in subjective value between a loss of \$1,200 and a loss of \$1,100. When the value functions for gains and for losses are pieced together, we obtain an S-shaped function of the type displayed in Figure 1.



Figure 1 A Hypothetical Value Function

The value function shown in Figure 1 is (a) defined on gains and losses rather than on total wealth, (b) concave in the domain of gains and convex in the domain of losses, and (c) considerably steeper for losses than for gains. The last property, which we label *loss aversion*, expresses the intuition that a loss of \$X is more aversive than a gain of \$X is attractive. Loss aversion explains people's reluctance to bet on a fair coin for equal stakes: The attractiveness of the possible gain is not nearly sufficient to compensate for the aversiveness of the possible loss. For example, most respondents in a sample of undergraduates refused to stake \$10 on the toss of a coin if they stood to win less than \$30.

The assumption of risk aversion has played a central role in economic theory. However, just as the concavity of the value of gains entails risk aversion, the convexity of the value of losses entails risk seeking. Indeed, risk seeking in losses is a robust effect, particularly when the probabilities of loss are substantial. Consider, for example, a situation in which an individual is forced to choose between an 85% chance to lose \$1,000 (with a 15% chance to lose nothing) and a sure loss of \$800. A large majority of people express a preference for the gamble over the sure loss. This is a risk seeking choice because the expectation of the gamble (-\$850) is inferior to the expectation of the sure loss (-\$800). Risk seeking in the domain of losses has been confirmed by several investigators (Fishburn and Kochenberger 1979; Hershey and Schoemaker 1980; Payne, Laughhunn, and Crum 1980; Slovic, Fischhoff, and Lichtenstein 1982). It has also been observed with nonmonetary outcomes, such as hours of pain (Eraker and Sox 1981) and loss of human lives (Fischhoff 1983; Tversky 1977; Tversky and Kahneman 1981). Is it wrong to be risk averse in the domain of gains and risk seeking in the domain of losses? These preferences conform to compelling intuitions about the subjective value of gains and losses, and the presumption is that people should be entitled to their own values. However, we shall see that an S-shaped value function has implications that are normatively unacceptable.

()

APPENDIX B

To address the normative issue we turn from psychology to decision theory. Modern decision theory can be said to begin with the pioneering work of von Neumann and Morgenstern (1947), who laid down several qualitative principles, or axioms, that should govern the preferences of a rational decision maker. Their axioms included transitivity (if A is preferred to B and B is preferred to C, then A is preferred to C), and substitution (if A is preferred to B, then an even chance to get A or C is preferred to an even chance to get B or C), along with other conditions of a more technical nature. The normative and the descriptive status of the axioms of rational choice have been the subject of extensive discussions. In particular, there is convincing evidence that people do not always obey the substitution axiom, and considerable disagreement exists about the normative merit of this axiom (e.g., Allais and Hagen 1979). However, all analyses of rational choice incorporate two principles: dominance and invariance. Dominance demands that if prospect A is at least as good as prospect B in every respect and better than B in at least one respect, then A should be preferred to B. Invariance requires that the preference order between prospects should not depend on the manner in which they are described. In particular, two versions of a choice problem that are recognized to be equivalent when shown together should elicit the same preference even when shown separately. We now show that the requirement of invariance, however elementary and innocuous it may seem, cannot generally be satisfied.

FRAMING OF OUTCOMES

Risky prospects are characterized by their possible outcomes and by the probabilities of these outcomes. The same option, however, can be framed or described in different ways (Tversky and Kahneman 1981). For example, the possible outcomes of a gamble can be framed either as gains and losses relative to the status quo or as asset positions that incorporate initial wealth. Invariance requires that such changes in the description of outcomes should not alter the preference order. The following pair of problems illustrates a violation of this requirement. The total number of respondents in each problem is denoted by *N*, and the percentage who chose each option is indicated in parentheses.

Problem 1 (N = 152): Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:

If Program A is adopted, 200 people will be saved. (72%)

If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved. (28%) Which of the two programs would you favor?

The formulation of Problem 1 implicitly adopts as a reference point a state of affairs in which the disease is allowed to take its toll of 600 lives. The outcomes of the programs include the reference state and two possible gains, measured by the number of lives saved. As expected, preferences are risk averse: A clear majority of respondents prefer saving 200 lives for sure over a gamble that offers a one-third chance of saving 600 lives. Now consider another problem in which the same cover story is followed by a different description of the prospects associated with the two programs:

Problem 2 (N = 155):

If Program C is adopted, 400 people will die. (22%)

If Program D is adopted, there is a one-third probability that nobody will die

and a two-thirds probability that 600 people will die. (78%)

It is easy to verify that options C and D in Problem 2 are undistinguishable in real terms from options A and B in Problem 1, respectively. The second version, however, assumes a reference state in which no one dies of the disease. The best outcome is the maintenance of this state and the alternatives are losses measured by the number of people that will die of the disease. People who evaluate options in these terms are expected to show a risk seeking preference for the gamble (option D) over the sure loss of 400 lives. Indeed, there is more risk seeking in the second version of the problem than there is risk aversion in the first.

The failure of invariance is both pervasive and robust. It is as common among sophisticated respondents as among naive ones, and it is not eliminated even when the same respondents answer both questions within a few minutes. Respondents confronted with their conflicting answers are typically puzzled. Even after rereading the problems, they still wish to be risk averse in the "lives saved" version; they wish to be risk seeking in the "lives lost" version; and they also wish to obey invariance and give consistent answers in the two versions. In their stubborn appeal, framing effects resemble perceptual illusions more than computational errors.

The following pair of problems elicits preferences that violate the dominance requirement of rational choice.

Problem 3 (N = 86): Choose between:

E. 25% chance to win \$240 and 75% chance to lose \$760 (0%)

F. 25% chance to win \$250 and 75% chance to lose \$750 (100%)

It is easy to see that F dominates E. Indeed, all respondents chose accordingly.

Problem 4 (N = 150): Imagine that you face the following pair of concurrent decisions. First examine both decisions, then indicate the options you prefer.

Decision (i) Choose between:

A. a sure gain of \$240 (84%)

B. 25% chance to gain \$1,000 and 75% chance to gain nothing (16%)

Decision (ii) Choose between:

C. a sure loss of \$750 (13%)

D. 75% chance to lose \$1,000 and 25% chance to lose nothing (87%)

As expected from the previous analysis, a large majority of subjects made a risk averse choice for the sure gain over the positive gamble in the first decision, and an even larger majority of subjects made a risk seeking choice for the gamble over the sure loss in the second decision. In fact, 73% of the respondents chose A and D and only 3% chose B and C. The same pattern of results was observed in a modified version of the problem, with reduced stakes, in which undergraduates selected gambles that they would actually play.

Because the subjects considered the two decisions in Problem 4 simultaneously, they expressed in effect a preference for A and D over B and C. The preferred conjunction, however, is actually dominated by the rejected one. Adding the sure gain of \$240 (option A) to option D

yields a 25% chance to win \$240 and a 75% chance to lose \$760. This is precisely option E in Problem 3. Similarly, adding the sure loss of \$750 (option C) to option B yields a 25% chance to win \$250 and a 75% chance to lose \$750. This is precisely option F in Problem 3. Thus, the susceptibility to framing and the S-shaped value function produce a violation of dominance in a set of concurrent decisions.

The moral of these results is disturbing: Invariance is normatively essential, intuitively compelling, and psychologically unfeasible. Indeed, we conceive only two ways of guaranteeing invariance. The first is to adopt a procedure that will transform equivalent versions of any problem into the same canonical representation. This is the rationale for the standard admonition to students of business, that they should consider each decision problem in terms of total assets rather than in terms of gains or losses (Schlaifer 1959). Such a representation would avoid the violations of invariance illustrated in the previous problems, but the advice is easier to give than to follow. Except in the context of possible ruin, it is more natural to consider financial outcomes as gains and losses rather than as states of wealth. Furthermore, a canonical representation (e.g., Problem 4) that exceeds the capabilities of intuitive computation even in simple problems. Achieving a canonical representation is even more difficult in other contexts such as safety, health, or quality of life. Should we advise people to evaluate the consequence of a public health policy (e.g., Problems 1 and 2) in terms of overall mortality, mortality due to diseases, or the number of deaths associated with the particular disease under study?

Another approach that could guarantee invariance is the evaluation of options in terms of their actuarial rather than their psychological consequences. The actuarial criterion has some appeal in the context of human lives, but it is clearly inadequate for financial choices, as has been generally recognized at least since Bernoulli, and it is entirely inapplicable to outcomes that lack an objective metric. We conclude that frame invariance cannot be expected to hold and that a sense of confidence in a particular choice does not ensure that the same choice would be made in another frame. It is therefore good practice to test the robustness of preferences by deliberate attempts to frame a decision problem in more than one way (Fischhoff, Slovic, and Lichtenstein 1980).

THE PSYCHOPHYSICS OF CHANCES

Our discussion so far has assumed a Bernoullian expectation rule according to which the value, or utility, of an uncertain prospect is obtained by adding the utilities of the possible outcomes, each weighted by its probability. To examine this assumption, let us again consult psychophysical intuitions. Setting the value of the status quo at zero, imagine a cash gift, say of \$300, and assign it a value of one. Now imagine that you are only given a ticket to a lottery that has a single prize of \$300. How does the value of the ticket vary as a function of the probability of winning the prize? Barring utility for gambling, the value of such a prospect must vary between zero (when the chance of winning is nil) and one (when winning \$300 is a certainty).

Intuition suggests that the value of the ticket is not a linear function of the probability of winning, as entailed by the expectation rule. In particular, an increase from 0% to 5% appears to have a larger effect than an increase from 30% to 35%, which also appears smaller than an increase from 95% to 100%. These considerations suggest a category-boundary effect: A change from impossibility to possibility or from possibility to certainty has a bigger impact than a comparable change in the middle of the scale. This hypothesis is incorporated into the curve dis-

played in Figure 2, which plots the weight attached to an event as a function of its stated numerical probability. The most salient feature of Figure 2 is that decision weights are regressive with respect to stated probabilities. Except near the endpoints, an increase of .05 in the probability of winning increases the value of the prospect by less than 5% of the value of the prize. We next investigate the implications of these psychophysical hypotheses for preferences among risky options.



In Figure 2, decision weights are lower than the corresponding probabilities over most of the range. Underweighting of moderate and high probabilities relative to sure things contributes to risk aversion in gains by reducing the attractiveness of positive gambles. The same effect also contributes to risk seeking in losses by attenuating the aversiveness of negative gambles. Low probabilities, however, are overweighted, and very low probabilities are either overweighted quite grossly or neglected altogether, making the decision weights highly unstable in that region. The overweighting of low probabilities reverses the pattern described above: It enhances the value of long shots and amplifies the aversiveness of a small chance of a severe loss. Consequently, people are often risk seeking in dealing with improbable gains and risk averse in dealing with unlikely losses. Thus, the characteristics of decision weights contribute to the attractiveness of both lottery tickets and insurance policies.

The nonlinearity of decision weights inevitably leads to violations of invariance, as illustrated in the following pair of problems:

Problem 5 (N = 85): Consider the following two-stage game. In the first stage, there is a 75% chance to end the game without winning anything and a 25% chance to move into the second stage. If you reach the second stage you have a choice between:

A. a sure win of \$30 (74%)

B. 80% chance to win \$45 (26%)

Your choice must be made before the game starts, i.e., before the outcome of the first stage is known. Please indicate the option you prefer.

Problem 6 (N = 81): Which of the following options do you prefer? C. 25% chance to win \$30 (42%)

D. 20% chance to win \$45 (58%)

Because there is one chance in four to move into the second stage in Problem 5, prospect A offers a .25 probability of winning \$30, and prospect B offers $.25 \times .80 = .20$ probability of winning \$45. Problems 5 and 6 are therefore identical in terms of probabilities and outcomes. However, the preferences are not the same in the two versions: A clear majority favors the higher chance to win the smaller amount in Problem 5, whereas the majority goes the other way in Problem 6. This violation of invariance has been confirmed with both real and hypothetical monetary payoffs (the present results are with real money), with human lives as outcomes, and with a nonsequential representation of the chance process.

We attribute the failure of invariance to the interaction of two factors: the framing of probabilities and the nonlinearity of decision weights. More specifically, we propose that in Problem 5 people ignore the first phase, which yields the same outcome regardless of the decision that is made, and focus their attention on what happens if they do reach the second stage of the game. In that case, of course, they face a sure gain if they choose option A and an 80% chance of winning if they prefer to gamble. Indeed, people's choices in the sequential version are practically identical to the choices they make between a sure gain of \$30 and an 85% chance to win \$45. Because a sure thing is overweighted in comparison with events of moderate or high probability (see figure 2), the option that may lead to a gain of \$30 is more attractive in the sequential version. We call this phenomenon the pseudo-certainty effect because an event that is actually uncertain is weighted as if it were certain.

A closely related phenomenon can be demonstrated at the low end of the probability range. Suppose you are undecided whether or not to purchase earthquake insurance because the premium is quite high. As you hesitate, your friendly insurance agent comes forth with an alternative offer: "For half the regular premium you can be fully covered if the quake occurs on an odd day of the month. This is a good deal because for half the price you are covered for more than half the days." Why do most people find such probabilistic insurance distinctly unattractive? Figure 2 suggests an answer. Starting anywhere in the region of low probabilities, the impact on the decision weight of a reduction of probability from *p* to p/2 is considerably smaller than the effect of a reduction from p/2 to 0. Reducing the risk by half, then, is not worth half the premium.

The aversion to probabilistic insurance is significant for three reasons. First, it undermines the classical explanation of insurance in terms of a concave utility function. According to expected utility theory, probabilistic insurance should be definitely preferred to normal insurance when the latter is just acceptable (see Kahneman and Tversky 1979). Second, probabilistic insurance represents many forms of protective action, such as having a medical checkup, buying new tires, or installing a burglar alarm system. Such actions typically reduce the probability of some hazard without eliminating it altogether. Third, the acceptability of insurance can be manipulated by the framing of the contingencies. An insurance policy that covers fire but not flood, for example, could be evaluated either as full protection against a specific risk (e.g., fire), or as a reduction in the overall probability of property loss. Figure 2 suggests that people greatly undervalue a reduction in the probability of a hazard in comparison to the complete elimination of that hazard. Hence, insurance should appear more attractive when it is framed as the elimination of risk than when it is described as a reduction of risk. Indeed, Slovic, Fischhoff,

105838_01_1-484_r5sk.indd

and Lichtenstein (1982) showed that a hypothetical vaccine that reduces the probability of contracting a disease from 20% to 10% is less attractive if it is described as effective in half of the cases than if it is presented as fully effective against one of two exclusive and equally probable virus strains that produce identical symptoms.

FORMULATION EFFECTS

So far we have discussed framing as a tool to demonstrate failures of invariance. We now turn attention to the processes that control the framing of outcomes and events. The public health problem illustrates a formulation effect in which a change of wording from "lives saved" to "lives lost" induced a marked shift of preference from risk aversion to risk seeking. Evidently, the subjects adopted the descriptions of the outcomes as given in the question and evaluated the outcomes accordingly as gains or losses. Another formulation effect was reported by McNeil, Pauker, Sox, and Tversky (1982). They found that preferences of physicians and patients between hypothetical therapies for lung cancer varied markedly when their probable outcomes were described in terms of mortality or survival. Surgery, unlike radiation therapy, entails a risk of death during treatment. As a consequence, the surgery option was relatively less attractive when the statistics of treatment outcomes were described in terms of mortality rather than in terms of survival.

A physician, and perhaps a presidential advisor as well, could influence the decision made by the patient or by the President, without distorting or suppressing information, merely by the framing of outcomes and contingencies. Formulation effects can occur fortuitously, without anyone being aware of the impact of the frame on the ultimate decision. They can also be exploited deliberately to manipulate the relative attractiveness of options. For example, Thaler (1980) noted that lobbyists for the credit card industry insisted that any price difference between cash and credit purchases be labeled a cash discount rather than a credit card surcharge. The two labels frame the price difference as a gain or as a loss by implicitly designating either the lower or the higher price as normal. Because losses loom larger than gains, consumers are less likely to accept a surcharge than to forgo a discount. As is to be expected, attempts to influence framing are common in the marketplace and in the political arena.

The evaluation of outcomes is susceptible to formulation effects because of the nonlinearity of the value function and the tendency of people to evaluate options in relation to the reference point that is suggested or implied by the statement of the problem. It is worthy of note that in other contexts people automatically transform equivalent messages into the same representation. Studies of language comprehension indicate that people quickly recode much of what they hear into an abstract representation that no longer distinguishes whether the idea was expressed in an active or in a passive form and no longer discriminates what was actually said from what was implied, presupposed, or implicated (Clark and Clark 1977). Unfortunately, the mental machinery that performs these operations silently and effortlessly is not adequate to perform the task of recoding the two versions of the public health problem or the mortality survival statistics into a common abstract form.

TRANSACTIONS AND TRADES

Our analysis of framing and of value can be extended to choices between multiattribute options, such as the acceptability of a transaction or a trade. We propose that, in order to evaluate

a multiattribute option, a person sets up a mental account that specifies the advantages and the disadvantages associated with the option, relative to a multiattribute reference state. The overall value of an option is given by the balance of its advantages and its disadvantages in relation to the reference state. Thus, an option is acceptable if the value of its advantages exceeds the value of its disadvantages. This analysis assumes psychological—but not physical—separability of advantages and disadvantages. The model does not constrain the manner in which separate attributes are combined to form overall measures of advantage and of disadvantage, but it imposes on these measures assumptions of concavity and of loss aversion.

Our analysis of mental accounting owes a large debt to the stimulating work of Richard Thaler (1980, 1985), who showed the relevance of this process to consumer behavior. The following problem, based on examples of Savage (1954) and Thaler (1980), introduces some of the rules that govern the construction of mental accounts and illustrates the extension of the concavity of value to the acceptability of transactions.

Problem 7: Imagine that you are about to purchase a jacket for \$125 and a calculator for \$15. The calculator salesman informs you that the calculator you wish to buy is on sale for \$10 at the other branch of the store, located 20 minutes' drive away. Would you make a trip to the other store?

This problem is concerned with the acceptability of an option that combines a disadvantage of inconvenience with a financial advantage that can be framed as a minimal, topical, or comprehensive account. The minimal account includes only the differences between the two options and disregards the features that they share. In the minimal account, the advantage associated with driving to the other store is framed as a gain of \$5. A topical account relates the consequences of possible choices to a reference level that is determined by the context within which the decision arises. In the preceding problem, the relevant topic is the purchase of the calculator, and the benefit of the trip is therefore framed as a reduction of the price, from \$15 to \$10. Because the potential saving is associated only with the calculator, the price of the jacket is not included in the topical account. The price of the jacket, as well as other expenses, could well be included in a more comprehensive account in which the saving would be evaluated in relation to, say, monthly expenses.

The formulation of the preceding problem appears neutral with respect to the adoption of a minimal, topical, or comprehensive account. We suggest, however, that people will spontaneously frame decisions in terms of topical accounts that, in the context of decision making, play a role analogous to that of "good forms" in perception and of basic-level categories in cognition. Topical organization, in conjunction with the concavity of value, entails that the willingness to travel to the other store for a saving of \$5 on a calculator should be inversely related to the price of the calculator and should be independent of the price of the jacket. To test this prediction, we constructed another version of the problem in which the prices of the two items were interchanged. The price of the calculator was given as \$125 in the first store and \$120 in the other branch, and the price of the jacket was set at \$15. As predicted, the proportions of respondents who said they would make the trip differed sharply in the two problems. The results showed that 68% of the respondents (N = 88) were willing to drive to the other branch to save \$5 on a \$15 calculator, but only 29% of 93 respondents were willing to make the same trip to save \$5 on a \$125 calculator. This finding supports the notion of topical organization of accounts, since the two versions are identical both in terms of a minimal and a comprehensive account.

The significance of topical accounts for consumer behavior is confirmed by the observation that the standard deviation of the prices that different stores in a city quote for the same product is roughly proportional to the average price of that product (Pratt, Wise, and Zeckhauser 1979). Since the dispersion of prices is surely controlled by shoppers' efforts to find the best buy, these results suggest that consumers hardly exert more effort to save \$15 on a \$150 purchase than to save \$5 on a \$50 purchase.

The topical organization of mental accounts leads people to evaluate gains and losses in relative rather than in absolute terms, resulting in large variations in the rate at which money is exchanged for other things, such as the number of phone calls made to find a good buy or the willingness to drive a long distance to get one. Most consumers will find it easier to buy a car stereo system or a Persian rug, respectively, in the context of buying a car or a house than separately. These observations, of course, run counter to the standard rational theory of consumer behavior, which assumes invariance and does not recognize the effects of mental accounting.

The following problems illustrate another example of mental accounting in which the posting of a cost to an account is controlled by topical organization:

Problem 8 (N = 200): Imagine that you have decided to see a play and paid the admission price of \$10 per ticket. As you enter the theater, you discover that you have lost the ticket. The seat was not marked, and the ticket cannot be recovered.

Would you pay \$10 for another ticket? Yes (46%) No (54%)

Problem 9 (N = 183): Imagine that you have decided to see a play where admission is \$10 per ticket. As you enter the theater, you discover that you have lost a \$10 bill.

Would you still pay \$10 for a ticket for the play? Yes (88%) No (12%)

The difference between the responses to the two problems is intriguing. Why are so many people unwilling to spend \$10 after having lost a ticket, if they would readily spend that sum after losing an equivalent amount of cash? We attribute the difference to the topical organization of mental accounts. Going to the theater is normally viewed as a transaction in which the cost of the ticket is exchanged for the experience of seeing the play. Buying a second ticket increases the cost of seeing the play to a level that many respondents apparently find unacceptable. In contrast, the loss of the cash is not posted to the account of the play, and it affects the purchase of a ticket only by making the individual feel slightly less affluent.

An interesting effect was observed when the two versions of the problem were presented to the same subjects. The willingness to replace a lost ticket increased significantly when that problem followed the lost-cash version. In contrast, the willingness to buy a ticket after losing cash was not affected by prior presentation of the other problem. The juxtaposition of the two problems apparently enabled the subjects to realize that it makes sense to think of the lost ticket as lost cash, but not vice versa.

The normative status of the effects of mental accounting is questionable. Unlike earlier examples, such as the public health problem, in which the two versions differed only in form, it can be argued that the alternative versions of the calculator and ticket problems differ also in substance. In particular, it may be more pleasurable to save \$5 on a \$15 purchase than on a

()

APPENDIX B

larger purchase, and it may be more annoying to pay twice for the same ticket than to lose \$10 in cash. Regret, frustration, and self-satisfaction can also be affected by framing (Kahneman and Tversky 1982). If such secondary consequences are considered legitimate, then the observed preferences do not violate the criterion of invariance and cannot readily be ruled out as inconsistent or erroneous. On the other hand, secondary consequences may change upon reflection. The satisfaction of saving \$5 on a \$15 item can be marred if the consumer discovers that she would not have exerted the same effort to save \$10 on a \$200 purchase. We do not wish to recommend that any two decision problems that have the same primary consequences should be resolved in the same way. We propose, however, that systematic examination of alternative framings offers a useful reflective device that can help decision makers assess the values that should be attached to the primary and secondary consequences of their choices.

LOSSES AND COSTS

Many decision problems take the form of a choice between retaining the status quo and accepting an alternative to it, which is advantageous in some respects and disadvantageous in others. The analysis of value that was applied earlier to unidimensional risky prospects can be extended to this case by assuming that the status quo defines the reference level for all attributes. The advantages of alternative options will then be evaluated as gains and their disadvantages as losses. Because losses loom larger than gains, the decision maker will be biased in favor of retaining the status quo.

Thaler (1980) coined the term "endowment effect" to describe the reluctance of people to part from assets that belong to their endowment. When it is more painful to give up an asset than it is pleasurable to obtain it, buying prices will be significantly lower than selling prices. That is, the highest price that an individual will pay to acquire an asset will be smaller than the minimal compensation that would induce the same individual to give up that asset, once acquired. Thaler discussed some examples of the endowment effect in the behavior of consumers and entrepreneurs. Several studies have reported substantial discrepancies between buying and selling prices in both hypothetical and real transactions (Gregory 1983; Hammack and Brown 1974; Knetsch and Sinden 1984). These results have been presented as challenges to standard economic theory, in which buying and selling prices coincide except for transaction costs and effects of wealth. We also observed reluctance to trade in a study of choices between hypothetical jobs that differed in weekly salary (S) and in the temperature (T) of the workplace. Our respondents were asked to imagine that they held a particular position (S_1, T_1) and were offered the option of moving to a different position (S_2, T_2) , which was better in one respect and worse in another. We found that most subjects who were assigned to (S_1, T_1) did not wish to move to (S_{2}, T_{2}) , and that most subjects who were assigned to the latter position did not wish to move to the former. Evidently, the same difference in pay or in working conditions looms larger as a disadvantage than as an advantage.

In general, loss aversion favors stability over change. Imagine two hedonically identical twins who find two alternative environments equally attractive. Imagine further that by force of circumstance the twins are separated and placed in the two environments. As soon as they adopt their new states as reference points and evaluate the advantages and disadvantages of each other's environments accordingly, the twins will no longer be indifferent between the two states, and both will prefer to stay where they happen to be. Thus, the instability of preferences

produces a preference for stability. In addition to favoring stability over change, the combination of adaptation and loss aversion provides limited protection against regret and envy by reducing the attractiveness of foregone alternatives and of others' endowments.

Loss aversion and the consequent endowment effect are unlikely to play a significant role in routine economic exchanges. The owner of a store, for example, does not experience money paid to suppliers as losses and money received from customers as gains. Instead, the merchant adds costs and revenues over some period of time and only evaluates the balance. Matching debits and credits are effectively canceled prior to evaluation. Payments made by consumers are also not evaluated as losses but as alternative purchases. In accord with standard economic analysis, money is naturally viewed as a proxy for the goods and services that it could buy. This mode of evaluation is made explicit when an individual has in mind a particular alternative, such as, "I can either buy a new camera or a new tent." In this analysis, a person will buy a camera if its subjective value exceeds the value of retaining the money it would cost.

There are cases in which a disadvantage can be framed either as a cost or as a loss. In particular, the purchase of insurance can also be framed as a choice between a sure loss and the risk of a greater loss. In such cases the cost-loss discrepancy can lead to failures of invariance. Consider, for example, the choice between a sure loss of \$50 and a 25% chance to lose \$200. Slovic, Fischhoff, and Lichtenstein (1982) reported that 80% of their subjects expressed a risk-seeking preference for the gamble over the sure loss. However, only 35% of subjects refused to pay \$50 for insurance against a 25% risk of losing \$200. Similar results were also reported by Schoemaker and Kunreuther (1979) and by Hershey and Schoemaker (1980). We suggest that the same amount of money that was framed as an uncompensated loss in the first problem was framed as the cost of protection in the second. The modal preference was reversed in the two problems because losses are more aversive than costs.

We have observed a similar effect in the positive domain, as illustrated by the following pair of problems:

Problem 10: Would you accept a gamble that offers a 10% chance to win \$95 and a 90% chance to lose \$5?

Problem 11: Would you pay \$5 to participate in a lottery that offers a 10% chance to win \$100 and a 90% chance to win nothing?

A total of 132 undergraduates answered the two questions, which were separated by a short filler problem. The order of the questions was reversed for half the respondents. Although it is easily confirmed that the two problems offer objectively identical options, 55 of the respondents expressed different preferences in the two versions. Among them, 42 rejected the gamble in Problem 10 but accepted the equivalent lottery in Problem 11. The effectiveness of this seemingly inconsequential manipulation illustrates both the cost-loss discrepancy and the power of framing. Thinking of the \$5 as a payment makes the venture more acceptable than thinking of the same amount as a loss.

The preceding analysis implies that an individual's subjective state can be improved by framing negative outcomes as costs rather than as losses. The possibility of such psychological manipulations may explain a paradoxical form of behavior that could be labeled the dead-loss effect. Thaler (1980) discussed the example of a man who develops tennis elbow soon after pay-

ing the membership fee in a tennis club and continues to play in agony to avoid wasting his investment. Assuming that the individual would not play if he had not paid the membership fee, the question arises: How can playing in agony improve the individual's lot? Playing in pain, we suggest, maintains the evaluation of the membership fee as a cost. If the individual were to stop playing, he would be forced to recognize the fee as a dead loss, which may be more aversive than playing in pain.

CONCLUDING REMARKS

The concepts of utility and value are commonly used in two distinct senses: (a) experience value, the degree of pleasure or pain, satisfaction or anguish in the actual experience of an outcome; and (b) decision value, the contribution of an anticipated outcome to the overall attractiveness or aversiveness of an option in a choice. The distinction is rarely explicit in decision theory because it is tacitly assumed that decision values and experience values coincide. This assumption is part of the conception of an idealized decision maker who is able to predict future experiences with perfect accuracy and evaluate options accordingly. For ordinary decision makers, however, the correspondence of decision values between experience values is far from perfect (March 1978). Some factors that affect experience are not easily anticipated, and some factors that affect decisions do not have a comparable impact on the experience of outcomes.

In contrast to the large amount of research on decision making, there has been relatively little systematic exploration of the psychophysics that relate hedonic experience to objective states. The most basic problem of hedonic psychophysics is the determination of the level of adaptation or aspiration that separates positive from negative outcomes. The hedonic reference point is largely determined by the objective status quo, but it is also affected by expectations and social comparisons. An objective improvement can be experienced as a loss, for example, when an employee receives a smaller raise than everyone else in the office. The experience of pleasure or pain associated with a change of state is also critically dependent on the dynamics of hedonic adaptation. Brickman and Campbell's (1971) concept of the hedonic treadmill suggests the radical hypothesis that rapid adaptation will cause the effects of any objective improvement to be short-lived. The complexity and subtlety of hedonic experience make it difficult for the decision maker to anticipate the actual experience that outcomes will produce. Many a person who ordered a meal when ravenously hungry has admitted to a big mistake when the fifth course arrived on the table. The common mismatch of decision values and experience values introduces an additional element of uncertainty in many decision problems.

The prevalence of framing effects and violations of invariance further complicates the relation between decision values and experience values. The framing of outcomes often induces decision values that have no counterpart in actual experience. For example, the framing of outcomes of therapies for lung cancer in terms of mortality or survival is unlikely to affect experience, although it can have a pronounced influence on choice. In other cases, however, the framing of decisions affects not only decision but experience as well. For example, the framing of an expenditure as an uncompensated loss or as the price of insurance can probably influence the experience of that outcome. In such cases, the evaluation of outcomes in the context of decisions not only anticipates experience but also molds it.

REFERENCES

- Allais, M., and O. Hagen, eds. 1979. *Expected Utility Hypotheses and the Allais Paradox*. Hingham, MA: D. Reidel.
- Bernoulli, D. 1954 [1738]. "Exposition of a New Theory on the Measurement of Risk." *Econometrica* 22:23–36.
- Brickman, P., and D. T. Campbell. 1971. "Hedonic Relativism and Planning the Good Society." In Adaptation Level Theory: A Symposium, ed. M. H. Appley. New York: Academic Press, 287–302.

Clark, H. H., and E. V. Clark. 1977. Psychology and Language. New York: Harcourt.

- Erakar, S. E., and H. C. Sox. 1981. "Assessment of Patients' Preferences for Therapeutic Outcomes." *Medical Decision Making* 1:29–39.
- Fischhoff, B. 1983. "Predicting Frames." *Journal of Experimental Psychology: Learning, Memory and Cognition* 9:103–16.
- Fischhoff, B., P. Slovic, and S. Lichtenstein. 1980. "Knowing What You Want: Measuring Labile Values." In *Cognitive Processes in Choice and Decision Behavior*, ed. T. Wallsten. Hillsdale, NJ: Erlbaum, 117–41.
- Fishburn, P. C., and G. A. Kochenberger. 1979. "Two-Piece von Neumann–Morgenstern Utility Functions." *Decision Sciences* 10:503–18.
- Gregory, R. 1983. "Measures of Consumer's Surplus: Reasons for the Disparity in Observed Values." Unpublished manuscript, Keene State College, Keene, NH.
- Hammack, J., and G. M. Brown Jr. 1974. *Waterfowl and Wetlands: Toward Bioeconomic Analysis*. Baltimore: Johns Hopkins University Press.
- Hershey, J. C., and P. J. H. Schoemaker. 1980. "Risk Taking and Problem Context in the Domain of Losses: An Expected-Utility Analysis." *Journal of Risk and Insurance* 47:111–32.
- Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47:263–91.
- ———. 1982. "The Simulation Heuristic." In *Judgment Under Uncertainty: Heuristics and Biases*, ed. D. Kahneman, P. Slovic, and A. Tversky. New York: Cambridge University Press, 201–08.
- Knetsch, J., and J. Sinden. 1984. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Quarterly Journal of Economics* 99:507–21.
- March, J. G. 1978. "Bounded Rationality, Ambiguity, and the Engineering of Choice." *Bell Journal of Economics* 9:587–608.
- McNeil, B., S. Pauker, H. Sox Jr., and A. Tversky. 1982. "On the Elicitation of Preferences for Alternative Therapies." *New England Journal of Medicine* 306:1259–62.
- Payne, J. W., D. J. Laughhunn, and R. Crum. 1980. "Translation of Gambles and Aspiration Level Effects in Risky Choice Behavior." *Management Science* 26:1039–60.
- Pratt, J. W., D. Wise, and R. Zeckhauser. 1979. "Price Differences in Almost Competitive Markets." *Quarterly Journal of Economics* 93:189–211.
- Savage, L. J. 1954. The Foundation of Statistics. New York: Wiley.
- Schlaifer, R. 1959. Probability and Statistics for Business Decisions. New York: McGraw-Hill.
- Schoemaker, P.J.H., and H. C. Kunreuther. 1979. "An Experimental Study of Insurance Decisions." *Journal of Risk and Insurance* 46:603–18.
- Slovic, P., B. Fischhoff, and S. Lichtenstein. 1982. "Response Mode, Framing, and Information-

۲

Processing Effects in Risk Assessment." In New Directions for Methodology of Social and Behavioral Science: Question Framing and Response Consistency, ed. R. Hogarth. San Francisco: Jossey-Bass, 21–36.
Thaler, R. 1980. "Toward a Positive Theory of Consumer Choice." <i>Journal of Economic Behavior</i>
 <i>and Organization</i> 1:39–60. . 1985. "Using Mental Accounting in a Theory of Consumer Behavior." <i>Marketing Science</i>
4:199–214. Tversky, A. 1977. "On the Elicitation of Preferences: Descriptive and Prescriptive Consider-
ations." In <i>Conflicting Objectives in Decisions</i> , ed. D. Bell, R. L. Kenney, and H. Raiffa. New York: Wiley, 209–22.

Tversky, A., and D. Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211:453–58.

von Neumann, J., and O. Morgenstern. 1947. *Theory of Games and Economic Behavior*, 2nd ed. Princeton: Princeton University Press.

-

۲

۲