JUDGMENT AND DECISION MAKING

B. A. Mellers¹, A. Schwartz², and A. D. J. Cooke³

¹Department of Psychology, Ohio State University, Columbus, Ohio 43210, e-mail: mellers.1@osu.edu; ²Department of Medical Education, University of Illinois, Chicago, Illinois 60612-7309; ³Marketing Department, University of Florida, Gainesville, Florida 32611

KEY WORDS: risk, emotions, choices, beliefs, utilities

ABSTRACT

For many decades, research in judgment and decision making has examined behavioral violations of rational choice theory. In that framework, rationality is expressed as a single correct decision shared by experimenters and subjects that satisfies internal coherence within a set of preferences and beliefs. Outside of psychology, social scientists are now debating the need to modify rational choice theory with behavioral assumptions. Within psychology, researchers are debating assumptions about errors for many different definitions of rationality. Alternative frameworks are being proposed. These frameworks view decisions as more reasonable and adaptive than previously thought. For example, “rule following.” Rule following, which occurs when a rule or norm is applied to a situation, often minimizes effort and provides satisfying solutions that are “good enough,” though not necessarily the best. When rules are ambiguous, people look for reasons to guide their decisions. They may also let their emotions take charge. This chapter presents recent research on judgment and decision making from traditional and alternative frameworks.

CONTENTS

INTRODUCTION ................................................................. 448
Rationality is a Single Correct Response ............................... 449
Rationality is Internal Coherence and Logical Consistency .......... 449
Rationality is the Same for Subjects and Experimenters ............ 450
Alternatives to Rational Choice Theory ............................... 450
THE DECISION MAKER ....................................................... 451

0066-4308/98/0201-0447$08.00
INTRODUCTION


This review examines assumptions about rationality and alternative frameworks. For many years, rational choice theory has been the dominant framework in economics, political science, finance, marketing, and other fields. Many scholars believed that failures of rationality could not survive competitive market forces. Violations were viewed as relatively trivial or artifactual; decision makers either learn quickly or they are eliminated from the game. Re-
search in judgment and decision making has demonstrated increasingly more violations of rational choice theory, and the importance of behavioral assumptions is now a lively topic of debate among social scientists. Behavioral assumptions are showing up in new subdisciplines. In addition to behavioral decision theory (Edwards 1961), we now have behavioral game theory (Camerer 1990), behavioral finance theory (Thaler 1993), and behavioral accounts of law (Sunstein 1997). These areas have identified both real-world and laboratory situations in which people violate fundamental precepts of rational choice.

In the meantime, psychologists in judgment and decision making have learned a great deal about errors and biases. Unlike economists, psychologists never questioned the need for behavioral assumptions. Instead, they have been questioning definitions of errors and biases within rational choice theory and are examining alternative frameworks. Errors are always defined relative to a normative framework that makes assumptions about human goals. Within the rational choice theory, definitions of errors are usually based on three faulty assumptions. First, there is a single correct response. Second, correct responses can be expressed as coherence within a system of preferences and beliefs. Third, the subject and the experimenter agree on what constitutes a correct response. Below we examine each assumption in detail.

**Rationality is a Single Correct Response**

During the last several years, researchers have pointed out that laboratory tasks often have many correct answers. Why? Tasks are inadequately specified, and different normative frameworks can apply. For example, if a Bayesian and a frequentist were asked, “What is the probability of discovering extraterrestrial life somewhere in the solar system in the next twenty years?” they would give very different, albeit “correct,” answers. The frequentist would balk at the question, while the Bayesian would provide a number between 0 and 1 that expresses a degree of belief. Even when theoretical positions are recognized, problems may lack sufficient detail for a single correct response. Birnbaum (1983) shows that the famous cab problem has many correct answers, depending on one’s theory of the witness. The Monty Hall problem—should a contestant on The Monty Hall show change doors after one door has been opened—has multiple solutions, depending on one’s assumptions about Monty. Even insensitivity to regression may be appropriate if people are neither taught nor told to minimize the sum of squared errors when making intuitive predictions.

**Rationality is Internal Coherence and Logical Consistency**

In the rational choice framework, rationality is expressed as internal coherence and logical consistency within a system of beliefs and preferences. This as-
sumption has been widely criticized. Gigerenzer (1991, 1996) points out that we are really interested in good judgment, and good judgment requires an analysis of content, in addition to laws, principles, and axioms. Gigerenzer (1996) and Cosmides & Tooby (1996) argue that good judgment is domain specific and should reflect basic principles of survival and adaptation. Kahneman (1994) suggests that logical analyses should be supplemented with substantive evaluations that assess the quality of decision outcomes. Are the person’s beliefs grossly out of kilter with available evidence? Does the decision damage the person’s interest? Finally, Hammond (1996) believes that people struggle with both coherence, or internal consistency of decisions, and correspondence, or empirical accuracy of decisions based on multiple fallible indicators. He argues for the integration of these perspectives in views of rationality.

Rationality is the Same for Subjects and Experimenters

Subjects and experimenters are typically assumed to agree on what constitutes rationality. This assumption has also been questioned (e.g. see Frisch & Clemen 1994). Many tasks have flat maxima or multiple good solutions (von Winterfeldt & Edwards 1986). Furthermore, thinking about one decision takes energy away from another. Subjects often have other concerns, not known to the experimenter.

Early metaphors for decision makers posited human beings as intuitive scientists, statisticians, and economists. Researchers investigated how well people sized up against professional standards of competence based on expected utility theory, Bayesian inference, and least squares regression. Some have argued that decision errors associated with these metaphors may not necessarily be errors within other ontological frameworks. Depending on the situation, people may be better understood as intuitive politicians who balance pressures from competing constituencies, intuitive prosecutors who demand accountability, or intuitive theologians who protect sacred values from contamination (Tetlock 1992).

Alternatives to Rational Choice Theory

One of the most interesting developments in the last five years has been a movement away from normative frameworks. March (1994) offers a dichotomy between preference-based choices and rule following. The traditional view is preference based; choices are consequential, and options are evaluated using prior assessments of beliefs and values. Rule following involves the application of rules or principles to situations. Elster (1989) refers to these decisions as norms or conventions. Although rule following has not received the same formal treatment as preference-based decisions, it appears to describe many automatic decisions that do not involve tradeoffs.
This review discusses both rule-based choices and preference-based choices involving tradeoffs between beliefs and values. The first and second sections begin with research on the decision maker and the decision task, respectively. The third and fourth sections discuss rule following and preference-based choices, respectively. The last section concludes with a few thoughts about the state of the field and future directions. The review is selective, not exhaustive. The reader is urged to read Camerer (1995) for a more extensive review of experimental economics and Dawes (1997) for more extensive coverage of judgmental errors and biases.

THE DECISION MAKER

There are many dimensions along which decision makers vary; this section focuses on two: risk and emotions. Several books and reviews have appeared on risk (Fischhoff et al 1997, Schoemaker 1993, Shapira 1995, Yates 1992) and emotions (Landman 1993, Parducci 1995, Roese & Olson 1995). We begin with the distinction between risk perceptions—how risky we view objects, hazards, or technologies—and risk attitudes—how willing we are to accept risk.

Risk Perceptions

Two approaches have been used to study individual differences in risk perceptions. In the first, risk is a multidimensional construct with dimensions labeled as dread, lack of familiarity, and lack of controllability (Fischhoff et al 1981). Slovic (1996) argues that those who have less trust in governments, institutions, and authorities perceive risks of hazards or technologies as greater than those with more trust. For example, environmental hazards are perceived as riskier by women than by men. Environmental disasters are perceived as riskier by blacks than by whites. Well-educated, conservative, white men perceive environmental hazards as least risky (Flynn et al 1994, Slovic et al 1993). Finally, experts and nonexperts differ in their estimates of environmental disasters, particularly for low-probability risks (Gregory et al 1996, Peters & Slovic 1997). Experts are less willing than the public to generalize from animal studies to human beings about chemical causes of cancer (Kraus et al 1992). Furthermore, experts show large affiliation effects; chemical risks are often perceived as lower by toxicologists in industry than by toxicologists in academia (Kraus et al 1992).

How accurate are the perceptions of these groups? Slovic (1996) argues that this question is impossible to answer because there is no single, objective definition of risk. Risk is a social construct invented to cope with the dangers and uncertainties of life. For example, between 1950 and 1970, coal mines became less risky in terms of deaths from accidents per ton of coal, but riskier in terms
of accidents per employee (Wilson & Crouch 1982). Was coal mining riskier in 1950 or 1970? There is no right answer. Likewise, there is no single, objective definition of safety. For example, airline safety can be measured on many dimensions, including the percentage of flights ending in accidents relative to total number of flights and the percentage of traveler deaths relative to total number of travelers. But there is no single definition of safety.

The second approach to risk, reviewed by Yates & Stone (1992), examines the perceived riskiness of monetary gambles. Over a decade ago, Coombs & Lehner (1984) found that losses have greater impact than gains, an asymmetry well known in choice behavior (Kahneman & Tversky 1979). Coombs & Lehner described this asymmetry in risk judgments with a bilinear model, similar to subjective expected utility theory with sign-dependent utilities and probabilities. Luce & Weber (1986) proposed a theory of risk judgments called conjoint expected risk. Risk perceptions were described as a weighted combination of three probabilities (winning, losing, and receiving nothing), expected gains (each gain raised to a power) conditional on winning, and expected losses (each loss raised to a different power) conditional on losing.

More recently, Weber et al (1992) find that, holding probability constant, the effect of a given outcome on risk judgments decreases as the number of other outcomes in the gamble increases. This averaging effect cannot be explained by the previous models. Weber et al propose a relative weight averaging model with sign-dependent utilities and probabilities. Although each of the models describes risk judgments in some contexts, none of them gives a complete account, because changes in the stimulus context can alter decision strategies used to form risk perceptions (Mellers & Chang 1994). Simply by including certain gambles within the stimulus set, experimenters can get subjects to change their strategies for judging risk.

This approach to risk perception has identified some cultural differences. Bontempo et al (1997) asked students in Hong Kong, Taiwan, the Netherlands, and the United States to rate the riskiness of monetary gambles. Responses are well-described by the conjoint expected risk model. Parameters of the model differ for subjects from Western countries and those from Asian countries. Western subjects place greater weight on the probabilities of losses, and Asians place greater emphasis on the magnitudes of losses. For Westerners, perceived risk decreases as some of the outcomes in a gamble improve and become positive. For Asians, perceived risk is less influenced by whether any of the outcomes are positive. Risk is clearly a cultural construct.

**Risk Attitudes**

In economic theories, risk attitudes are measured by revealed preferences. Consider a choice between a gamble and a sure thing equal to the expected
value of the gamble. People who choose the sure thing are said to have risk-averse preferences, and those who choose the gamble have risk-seeking preferences. Preferences are often risk averse in the domain of gains. Kahneman & Lovallo (1993) point out that risk premiums (differences between the expected value of a gamble and its certainty equivalent) can be substantially reduced if risks are aggregated over time. Thaler et al (1997) provide additional support for this claim.

Although preferences are typically risk averse in the gain domain, they are frequently risk seeking in the loss domain, a result known as the reflection effect (Kahneman & Tversky 1979). In earlier research, the effect has been attributed to utility functions that differ for gains and losses. More recently, different weighting functions for gains and losses have been proposed instead of, or in addition to, changes in utilities. March (1996) examines whether preferences for risk can be described from experienced outcomes. Consider a two-alternative, forced-choice task with variable reinforcement. Reinforcement learning theories assume that choice depends only on the outcomes experienced. March defines a set of simple stochastic models that describe trial-by-trial learning and shows what happens when a learner is confronted with options of variable risk over many trials. When experienced outcomes are positive, learners favor less risky alternatives. When experienced outcomes are negative, learners favor riskier alternatives in the short run and risk neutrality in the long run. In short, the tendency for greater risk aversion with gains than with losses is predicted by simple theories of accumulated learning.

Do decision makers believe their own risk attitudes change across gain and loss domains? Weber & Milliman (1997) hypothesize that when risk preferences are defined by the decision maker, not by economic theory, perceived-risk attitudes will show greater consistency across domains. Weber & Milliman present subjects with pairs of gambles and measure both preferences for gambles and risk perceptions. The majority of subjects choose gambles perceived as less risky in both domains. Perceptions of risk vary across domains, but perceived-risk attitudes are more stable and consistent than risk attitudes defined by economic theory. Mellers et al (1997b) find similar results.

**Emotions**

Emotions have powerful effects on decisions. Moreover, the outcomes of decisions have powerful effects on emotions. This section reviews research on both predecision and postdecision affect.

**Predecision Affect**  Isen (1993) argues that positive emotions increase creative problem solving and facilitate the integration of information. Estrada et al (1994) find that doctors in whom positive affect has been induced inte-
grate information more efficiently than do controls, show less anchoring on earlier diagnoses, and display more creativity in their thinking. Positive feelings can promote variety seeking (Kahn & Isen 1993), overestimation of the likelihood of favorable events, and underestimation of the likelihood of unfavorable events (Nygren et al 1996, Wright & Bower 1992). In contrast, Bodenhausen et al (1994) find that people in positive moods are likely to engage in more stereotyped thinking than people in neutral moods. However, the effect vanishes when people are held accountable for their judgments.

Negative affect can produce a narrowing of attention and a failure to search for new alternatives (Fiedler 1988). People in negative moods make more attribute-based comparisons than alternative-based comparisons (Luce et al 1997). In addition, they make faster and less discriminate use of information that can increase choice accuracy in easier tasks and decrease it in harder tasks.

Research in this domain often treats emotions as a unidimensional construct, ranging from positive to negative. Lewinsohn & Mano (1993) propose a two-dimensional model of affect, based on pleasantness and arousal. People in pleasant moods deliberate longer, use more information, and reexamine more information than others. People in aroused states tend to take more risks. Those who are aroused and in unpleasant moods employ simpler decision strategies and form more polarized judgments (Mano 1992, 1994).

Even a two-dimensional model seems inadequate for describing emotional experiences. Anger, sadness, and disgust are all forms of negative affect, and arousal does not capture all of the differences among them. Furthermore, many emotions, such as parental love, are domain specific. A more detailed approach is required to understand relationships between emotions and decisions.

POSTDECISION AFFECT Most of us know all too well the feeling of regret that can follow a decision. Gilovich & Medvec (1994, 1995) show that in the short term, people feel greater regret about actions than inactions, but in the long term, people feel greater regret about inactions than actions. Gilovich & Medvec suggest that time reduces the sting of regrettable actions and increases the sadness of regrettable inactions. In contrast, Kahneman (1995) believes people regret actions more than inactions throughout their lives. He contends that Gilovich & Medvec are measuring two distinct emotions, one being an intense, “hot” feeling that accompanies action and the other being a reflective, “wistful” feeling that captures the sadness of missed opportunities.

Mellers et al (1997a) devised a paradigm for measuring both choices and affective responses to monetary outcomes of gambles. After a choice, subjects learn the outcome of the chosen gamble and describe their emotional response
to it on a scale ranging from very elated to very disappointed. This paradigm allows the estimation of decision utilities from choices and experienced utilities from emotions. Decision utilities differ from hedonic responses in two important respects. First, unlike decision utilities, experienced utilities are influenced by subjective probabilities. Surprising wins are more pleasurable than expected wins, and surprising losses are more disappointing. Second, unlike decision utilities, experienced utilities depend on counterfactual possibilities. Obtained outcomes are evaluated relative to what might have happened under different states of the world and different choices. These comparisons can make larger losses feel less painful than smaller losses and smaller gains feel more pleasurable than larger gains, a result also found by Boles & Messick (1995). Mellers et al provide an account of emotional responses that they call decision affect theory. With some additional assumptions, this theory can predict choices from emotions.

Memories of hedonic experiences can be important guides to future choice. Kahneman and his colleagues show that when we make global evaluations of past experiences, we are often insensitive to the duration of the experience (Fredrickson & Kahneman 1993, Varey & Kahneman 1992). In one study, Redelmeier & Kahneman (1996) examined moment-to-moment and retrospective evaluations of the pain experienced by patients undergoing diagnostic colonoscopy. Patients indicated their discomfort every 60 s during the procedure and their overall discomfort at the end. The duration of the procedure, which ranged from 4 min to 69 min, does not predict retrospective evaluations. Instead, a peak-end rule, representing an average of the worst moments and the final moments of the experience, predicts global hedonic responses. In other experiments, Kahneman and his colleagues show that by adding diminishing pain to the end of a painful experience, global evaluations can be made more positive (Kahneman et al 1993a). These results have both humane and Orwellian implications and suggest enormous possibilities for decision engineering.

THE DECISION TASK

Framing effects, stimulus contexts, environments, and response modes might seem innocuous, but they can profoundly shape decisions (Payne et al 1992). Preferences can reverse depending on each of these factors. These effects have important implications for policy making, market decisions, and pollsters.

Framing Effects

One of the most basic findings in decision making is what Slovic (1972) called the “concreteness principle”: people often accept and use information in the
form in which they receive it. When this principle holds, preferences for identical options with different reference points can reverse (Tversky & Kahneman 1981). These effects have been widely investigated (de Dreu et al 1994, Kashiina & Maher 1995, Paese et al 1993, Ritov et al 1993, Schweitzer 1995, Sullivan & Kida 1995). For example, Johnson et al (1993) show that preferences for insurance coverage can vary, depending on whether premiums are described as rebates or deductibles. At the time they were conducting their experiments, New Jersey and Pennsylvania offered lower insurance rates if drivers would give up the right to sue other drivers in a collision. In New Jersey, the default option did not include the right to sue, although the driver could purchase it at additional cost. In Pennsylvania, the default option included the right to sue, although the driver could decline it and receive a cost reduction. Prices for comparable coverage were roughly the same. When offered a choice between the two policies, only 20% of New Jersey drivers bought the right to sue, but 75% of Pennsylvania drivers purchased it. Johnson et al estimated that if limited tort had been the default in Pennsylvania, drivers would have saved approximately $200 million in insurance costs.

Not everyone who looks for framing effects finds them, and there are undoubtedly many reasons (Christensen et al 1995, Fagley & Miller 1990). First, framing effects are often examined with verbal scenarios, and small changes in wording can have big effects on preferences (Schneider 1992). In the Asian disease problem (Tversky & Kahneman 1981), one group of subjects choose between two programs designed to combat a disease that is expected to kill 600 people. If one program is adopted, 200 people will be saved, and if the other program is adopted, there is one third probability that 600 people will be saved and two thirds probability that no people will be saved. Another group of subjects choose between the programs described in terms of lives lost. If one program is adopted, 400 people will die, and if the other program is adopted, there is one third probability that nobody will die, and two thirds probability that 600 people will die. Kühberger (1995) notes that outcomes in the Asian disease problem are inadequately specified; knowing that 200 people will be saved does not tell us explicitly what will happen to the other 400 people. When he makes outcomes explicit, preference reversals vanish.

A second reason for the absence of framing effects is because researchers have manipulated the salience of the good or bad features of the outcomes rather than the reference point (van Schie & van der Pligt 1995). For example, descriptions of beef as 90% lean or 10% fat emphasize either the positive or negative features of the beef, respectively, but do not alter the reference point. van Schie & van der Pligt show that emphasis on positive features promotes greater risk-seeking preferences in both the gain and loss domains, and emphasis on negative features promotes greater risk aversion in both domains.
Third, the magnitude of framing effects may depend on the content domain. Frisch (1993) asked subjects to compare identical problems framed differently and decide whether the problems should be treated the same. Subjects tend to think that problems with well-specified monetary gains and losses are equivalent, but other problems, such as those involving sunk costs, seem different. When asked why, they say those problems have different emotional consequences.

**Stimulus Contexts and Environments**

Preferences are not created in a vacuum; they depend on the stimulus context. This context might include the environment or the local stimulus context in the world. It might also include a larger context based on the decision maker’s past and present experiences.

In inference tasks, people might be asked to consider which of two cities, San Francisco, California, or Columbus, Ohio, has a larger population. They search through their memories to make a response by recalling cues of varying validity. Gigerenzer (1997) and Gigerenzer & Goldstein (1996) examine a simple, lexicographic rule called “take the best, ignore the rest.” Cues are represented in memory as binary variables. The first cue is recognition (i.e. does the decision maker recognize the cities?). If only one city is recognized, that city is assumed to be larger. If neither city is recognized, a guess is made. If both cities are recognized, the search continues and the cue with the highest perceived validity is assessed. If one city has a value on that cue, that city is assumed to be larger. If neither city has a cue, the search continues. Likewise, if both cities have cues, the search continues. With relatively few cues, “take the best” can actually outperform linear regression. As the number of cues increases, linear regression outperforms “take the best” (Gigerenzer 1997). Predicted performance based on this simple, satisfying rule varies, depending on the number and validity of the cues in the environment.

In choice tasks, there is a local stimulus context created by the choice set and a global stimulus context provided by all of the choice sets presented for judgment. Asymmetric dominance is a well-known local context effect (Huber et al 1982). Suppose a decision maker chooses between two options, A and B. Later, the decision maker chooses among three options, A, B, and C. The new option, C, is dominated by B but not A. Huber et al find that the relative preference for B over A increases in the larger choice set, a violation of the property of independence of irrelevant alternatives. Several authors have suggested theories of this effect in which weights and/or utilities of the attributes change with the addition of new options (Ariely & Wallsten 1995, Simonson & Tversky 1992, Tversky & Simonson 1993, Wedell 1991, Wedell & Pettibone 1996).
The range and spacing of the attributes presented for choice determine the global stimulus context. Mellers & Cooke (1996) and Simonson & Tversky (1992) find that a given attribute difference has a greater effect on choice when the global range is narrow than when it is wide, and preferences between identical pairs of options can reverse in different global contexts. Such effects cannot be explained as response biases, because response transformations are typically assumed to be monotonic functions that could not produce ordinal changes in preferences.

What psychological processes are affected by the global stimulus context? Mellers & Cooke (1994) and Cooke & Mellers (1997) propose that changes in range and spacing influence the utilities of the attributes in a manner consistent with range-frequency theory (Parducci 1995). In contrast, Tversky & Simonson (1993) propose that changes in the global context affect attribute weights. An earlier account, proposed by Simonson (1989), posits that contextual effects influence the relative justifiability of response options. These theories are not mutually exclusive, and Wedell & Pettibone (1996) find support for both the range-frequency and justification accounts.

**WEIGHT ELICITATION** Weights are important in applied contexts. In decision analysis, options are typically decomposed into weights and scales, and information is combined by means of normative rules (Keeney & Raiffa 1976, von Winterfeldt & Edwards 1986). Weights should be sensitive to scale changes: If the range of an attribute presented for judgment is reduced by one half, the effective weight of that attribute should be doubled to compensate. Researchers have tested the sensitivity of different procedures to the range of attributes presented within an experiment. Direct weight assessments do not vary greatly with changes in attribute range (Fischer 1995, Mellers & Cooke 1994, von Nitzsch & Weber 1993). Comparative weight elicitation procedures, such as swing weights, ratio weights, and tradeoff weights, show greater range sensitivity, although less than what is required to compensate for the change in range (Fischer 1995, von Nitzsch & Weber 1993, Weber & Borcherding 1993).

Edwards & Barron (1994) suggest that by eliciting rank orders of importance over all attributes and using rank-ordered centroid weights, one has nearly the same accuracy as is found with more complex methods. Barron & Barrett (1996a) compare three methods for aggregating rank order weights: (a) divide each rank by the sum of the ranks, (b) divide the reciprocal of each rank by the sum of the reciprocals, and (c) average the vertices in the weight space to produce rank-ordered centroid weights. All three procedures work well, although rank-ordered centroid weights consistently outperform the others. Procedures are also given for sensitivity analyses (Barron & Barrett 1996b).
Response Mode Effects

Preferences have been shown to vary with the method of assessment. Consider two gambles, A and B, having equal expected values and different variances. People may assign a higher value to A than B in a pricing task but choose B over A in a choice task. Preference reversals are robust; they have been demonstrated with prices and choices, ratings and choices, and other pairs of response modes. Furthermore, when confronted with the reversals, most subjects adamantly defend them (Ordóñez et al 1995). This simple fact makes it extremely difficult to call preference reversals forms of decision errors.

Tversky et al (1988) propose a compatibility hypothesis to explain reversals: When attributes are compatible with the response scale, they are assigned greater weight. This hypothesis is offered for preference reversals between gambles using price and rating tasks. Mellers et al (1992a) propose an alternative account of preference reversals between gambles with the same two tasks called change-of-process theory. Decision strategies depend on the task, but utilities remain constant. They provide a test between change-of-process and contingent weighting and find evidence supporting the change-of-process theory.

There are undoubtedly multiple causes for preference reversals, depending on the options and the tasks. Fischer & Hawkins (1993) contrast compatibility of attributes and responses with compatibility of tasks and strategies. The latter type of compatibility implies that qualitative tasks, such as choice, induce qualitative strategies (e.g. lexicographic orders) favoring the more prominent attribute, and quantitative tasks, such as matching, evoke quantitative strategies (e.g. averaging) with more uniform weighting of attributes. Fischer & Hawkins find that preference reversals between riskless options using choices and ratings are better described by strategy compatibility than scale compatibility. That is, the more important attribute is weighted more heavily in choices than ratings, a result also found by Mellers et al (1995) and Mellers & Cooke (1996).

Others have speculated that it is not strategy compatibility but rather the salience of comparisons in choices versus ratings that produces preference reversals (Hsee 1996, Markman & Medin 1995). Nowlis & Simonson (1997) note that comparable attributes have greater influence in choice tasks. They ask subjects to evaluate products described by comparable attributes (e.g. price) and less-comparable attributes (e.g. brand name) and find that comparable attributes have greater effects in choices than ratings.

RULE FOLLOWING

We base many of our decisions on rules or heuristics that convey information about who we are and how we interact with others. These decisions may have
involved tradeoffs at one point, but they have become “generic” applications of rules to situations. They may express habits, such as when we wake up in the morning or which foods we purchase at the grocery store. They may convey a personal or moral identity, such as kindness or honesty. They may also convey a social identity, including professional or family ties. Regardless of the function of the rule, this procedure minimizes effort and allows us to turn our attention to other matters.

This section begins with a discussion of rule following for individual and social decisions. Sometimes rules do not apply, either because the rules are poorly defined or because they conflict with other rules. In these cases, decision makers may search for reasons or emotions to guide their choices.

Individual and Social Decisions

Prelec & Herrnstein (1991) discuss cases in which people avoid cost-benefit analysis and use prudential rules for moral considerations and matters of self-control. For example, what is the cost of eating one piece of chocolate cake or taking one car trip without a seat belt? Adopting a rule is especially useful for controlling one’s behavior when the impact of the act is felt only with repetition (e.g. smoking a cigarette), with a delay between costs and benefits (e.g. dieting), or with benefits that are hard to imagine (e.g. spending versus saving). Following a rule minimizes effort and allows people to avoid difficult tradeoffs.

Fiske (1992) maintains that social decisions can be described by four basic rules: communal sharing, authority ranking, equality matching, and market pricing. Communal sharing stresses common bonds among group members, as found with families, lovers, and nations. Authority ranking focuses on inherent asymmetries in relationships; some people have higher rank, privilege, or prestige than others. Equality matching stresses reciprocity. Examples include babysitting cooperatives or car pooling, where one should get back whatever one puts in. In market pricing, decisions are governed by supply and demand, expected utilities, or tradeoffs between costs and benefits.

What happens when the decision maker applies the wrong rule? Fiske & Tetlock (1997) examine such cases and, following Tetlock et al (1996), label them taboo tradeoffs. Proposals to place monetary values on things we think of as priceless, such as children, body organs, or votes, do not just trigger cognitive confusion—they activate moral outrage. Most people respond with contempt and wish to punish norm violators; taboo tradeoffs are threats to both personal identities and the social order.

Reason-Based Choice

When rules don’t single out one best action (Tversky & Shafir 1992a), people may look for reasons to make choices. Reasons may be lists of pros and cons,
or they may take the form of stories. Pennington & Hastie (1992, 1993) argue that jurors construct stories to explain the facts. Pennington & Hastie present evidence to subjects either as stories or issues and find that story organizations result in stronger and more confident jury decisions than issue organizations.

Reasons allow us to justify decisions to ourselves (Hogarth & Kunreuther 1995). Shafir et al (1993) create decisions in which people have reasons to choose an act $X$ if an event $A$ occurs, and different reasons to choose $X$ if $A$ does not occur. But if $A$ is unknown, there are no reasons to choose $X$, so it is rejected. This pattern violates Savage’s sure-thing principle. For example, Tversky & Shafir (1992b) ask subjects to assume they just took a tough qualifying exam. One group is told they passed, another group is told they failed, and a third group is told they will learn the results tomorrow. Each group is offered a choice among buying a vacation to Hawaii on sale today only, not buying the vacation, or paying $5 to retain the right to buy the vacation package tomorrow. The majority of those who think they passed or failed the exam select the vacation, but the majority of those who don’t know the results want to retain the right to buy the vacation tomorrow, presumably because they have no reason to purchase the package today.

Reasons also allow us to justify decisions to others. Justifiability can increase certain decision errors and decrease others (Simonson & Nye 1992, Tetlock & Kim 1987). Tetlock & Boettger (1994) examine situations in which people make choices between the status quo and options that provide gains for society as a whole, but impose losses on an identifiable minority. In these cases, accountability to others can enhance both loss aversion and decision avoidance. People who must justify past acts or expenses to others also get locked into decisions, a finding known as commitment to sunk costs. Arkes (1996) examines how people who are motivated to avoid the appearance of wastefulness compromise their self-interests and attend to sunk costs. Commitment to sunk costs has received much attention (Garland & Newport 1991, Larrick et al 1993, McCarthy et al 1993), and some have found methods for attenuating it (Simonson & Staw 1992, Tan & Yates 1995).

**Emotion-Based Choice**

Since the work of Bell (1982, 1985) and Loomes & Sugden (1982, 1986), there has been growing interest in the role of anticipated emotions in choice. Anticipated regret influences choices between gambles (Bar-Hillel & Neter 1996), medical decisions (Ritov & Baron 1990), consumer products (Simonson 1992), and sexual practices (Richard et al 1996). Baron and his colleagues demonstrate that anticipated regret can also produce an omission bias—a preference inaction over action (Baron 1994, Baron & Ritov 1994, Spranca et al 1991). For example, Ritov & Baron (1990) find that people prefer not to vacci-
nate their child when the vaccine has potentially fatal side effects, even if the death rate from the vaccine is a fraction of the death rate from the disease. People anticipate regret about causing their child’s death, but by avoiding the vaccine, they actually increase their child’s risk of dying.

Outcome feedback also influences anticipated regret. When people know they will learn the outcomes of unchosen options, they often make choices that minimize chances of feeling regret (Josephs et al 1992, Ritov & Baron 1995, Ritov 1996, Zeelenberg et al 1996). Tests of this theory have been done with pairs of gambles having equal expected values, and results have been consistent with the theory. However, support declines with pairs of gambles that differ in expected value (BA Mellers, A Schwartz & I Ritov, working paper).

PREFERENCE-BASED DECISIONS

Preference-based choices are consequential because actions depend on beliefs about the value of future outcomes. This section begins with a discussion of beliefs and values and then presents theories of choices and certainty equivalents. See Edwards (1992) for a more extensive review of utility theories.

Beliefs

Decisions are based on beliefs about the likelihood of future events. Those beliefs are expressed as probability judgments, judgments under uncertainty, and confidence judgments.

PROBABILITY JUDGMENTS Koehler (1996) examines research on base-rate neglect and argues that base rates are frequently used in decision making. The degree of use depends on task structure and representation. Within-subject designs, direct experience, frequentistic problems, unambiguous sample spaces, and random sampling promote base-rate usage (Cosmides & Tooby 1996, Gigerenzer & Hoffrage 1995, Tversky & Kahneman 1974, Weber et al 1993). Kruschke (1996) examines base rates in category learning tasks. He proposes that when subjects receive feedback, they quickly learn base rates. Common categories are encoded first, and rare categories are encoded later. As a result, people learn typical features of common categories and unusual features of rare categories. This theory predicts base-rate neglect (Gluck & Bower 1988) and the inverse base-rate effect in which people judge higher base-rate events as less likely than lower base-rate events (Medin & Edelson 1988, Nelson 1996).

Tversky & Koehler’s (1994) support theory offers an account of explicit and implicit disjunctive probability judgments. In support theory, subjective probability is assigned to hypotheses. Subjective probability increases as hypotheses are “unpacked” into more explicit disjunctions. Judged probabi-
ties are complementary in the binary case but subadditive in the general case. When people are asked, “What is the probability that a given person selected at random from a population will die from an accident?”; judgments are less than the sum of the judgments of the components (e.g. car crashes, plane crashes, fire, drowning, and other accidents). Support theory explains conjunction fallacies, hypothesis generation, decisions under uncertainty, and fault-tree errors (Russo & Kolzow 1994). It has also been used to describe choice under uncertainty (Fox & Tversky 1997).

**JUDGMENTS UNDER UNCERTAINTY** Decisions about financial investments, litigation, environmental disasters, and insurance are usually based on a lack of knowledge about relevant probabilities. In these situations, people are often ambiguity averse—they prefer known probability distributions over uncertain probability distributions. Actuaries suggest higher warranty prices for ambiguous probabilities than for well-specified probabilities (Hogarth & Kunreuther 1992), and underwriters set higher insurance premiums for ambiguous probabilities and losses than for well-specified probabilities and losses (Kunreuther et al 1995). See Camerer & Weber (1992) for a review of the literature on ambiguity.

Heath & Tversky (1990) note an exception to ambiguity aversion. People often prefer to bet on their own (ambiguous) beliefs over matched chance events when they feel competent about a knowledge domain. What makes them feel competent? Fox & Tversky (1995) suggest that feelings of competence occur when people have clear versus ambiguous knowledge. Fox & Tversky argue that the contrast between differential knowledge states occurs in all previous tests of ambiguity aversion. Subjects typically make choices between clear and ambiguous urns. In these comparative contexts, subjects will pay more for gambles based on clear than ambiguous probabilities. But in non-comparative contexts, subjects value the gambles equally. These results highlight the importance of local contexts in theories of ambiguity.

**CONFIDENCE JUDGMENTS** Many studies have examined internal uncertainty, such as the confidence that people place in their own abilities. Typical tasks involve one’s belief in one’s performance on general knowledge tests. Lack of calibration based on comparisons of confidence judgments against percentages of correct items have led researchers to argue that people are often overconfident. Recently, this conclusion has been criticized on two grounds.

First, some authors argue that overconfidence is simply regression to the mean. Erev et al (1994), Dawes & Mulford (1996), and Pfleifer (1994) show that when relative frequencies (or percentages of correct items) are averaged over confidence ratings, people are often overconfident. But when confidence ratings are averaged over relative frequencies, people are often underconfident or conservative. In short, the same set of data can show overconfidence and un-
derconfidence, depending on the analysis. Furthermore, the “hard-easy” effect—the fact that hard tests are often associated with overconfidence and easy tests with underconfidence—is also expected from a regression interpretation.

To avoid this criticism, some studies compare means, or average confidence ratings, with overall percentage correct. Average confidence ratings are sometimes greater than the average number of correct items, consistent with overconfidence (Brenner et al. 1996, Griffin & Tversky 1992, Schneider 1995). Other studies report no average difference (Juslin 1994). Reasons for the discrepancies are unclear.

Second, Gigerenzer et al. (1991) and Gigerenzer (1993) argue that overconfidence is not the result of judgment errors, because it only occurs when test items are sampled in nonrandom ways. When questions are selected randomly, overconfidence disappears (Juslin 1994). Using over 25 studies, Juslin et al. (1997) demonstrate that overconfidence is only mildly related to the hard-easy effect. However, Griffin & Tversky (1992) find that even with random sampling, overconfidence remains. Reasons for the discrepancies are unclear.

Gigerenzer et al. (1991) give another reason why overconfidence may not be a judgment error. They point out that if overconfidence is a general phenomena, it should appear regardless of whether it is measured with single test items or overall number of items. Results often show greater overconfidence with single items than judgments of percentage correct (Gigerenzer et al. 1991, Sniezek et al. 1990). Once again, there are conflicting results (Griffin & Tversky 1992), and reasons for the discrepancies remain unclear.

BAYESIAN BELIEF NETWORKS Inference, a process of deriving a conclusion from initial information and a set of rules, plays a key role in expert systems. The most common form of inference comes from Boolean logic in which statements are either true or false. Sometimes, however, evidence is neither true nor false; it is uncertain. In these cases, Bayesian belief networks can be extremely useful (Pearl 1988). Bayesian belief networks consist of nodes, representing probabilistic variables, and links, representing relations between nodes. Networks provide mechanisms for combining uncertain information and making probabilistic inferences. Decision problems that, in the past, were hopelessly complex and unmanageable, are made both visually simple and intuitively obvious largely due to assumptions about conditional independence.

Influence diagrams are used for making decisions. They contain value nodes, decision nodes, and Bayesian belief networks. Influence diagrams also provide power and visual simplicity (Howard & Matheson 1984). Much of the material published on Bayesian belief networks and influence diagrams is fairly technical. For introductory material on Bayesian belief networks, see Shafer (1996) and Morawski (1989). Edwards (1991) provides a paper on ap-
Applications in legal domains. More information on both theory and applications is available on the World Wide Web.

**Values**

The endowment effect refers to the observation that people value objects they own more than objects that are not part of their subjective endowment (Kahneman et al. 1990). Not only current ownership but history of ownership affects value (M. Strahilevitz & G.F. Loewenstein, manuscript in preparation). For objects in one’s possession, value increases with the duration of ownership. For objects that are currently not in one’s possession but were at one time, value increases with the duration of past ownership.

Loewenstein & Issacharoff (1994) further demonstrate that value is influenced by how the object was obtained. People who obtain an object due to exemplary performance value that object more highly than people who obtain the same object due either to chance or to poor performance. Their results have implications for public policies, such as housing programs. Policies that give homes away to lower income families may be less effective at improving neighborhoods than policies requiring families to purchase homes, even at extremely low prices. These results converge with those of Arkes et al. (1994) who find that windfall gains are spent more readily than other types of assets, presumably because they are valued less. Similarly, unexpected tax rebates, lottery winnings, and inheritances may have less value than earned income.

**DELAYED VALUES** When the outcomes of decisions are delayed in time, people often discount the value of the delayed outcome. Discounting functions are often assumed to be exponential, although many experiments suggest they follow a hyperbolic rather than exponential form. Hyperbolic functions imply that when faced with an inferior option now or a superior option later, people want the option now. However, their preferences reverse when the same two options are offered with a constant delay added to each. Loewenstein & Elster (1992) provide a review of this literature.

Risky and Uncertain Choice

Standard economic theories are based on the assumption that utilities and beliefs are separable, but there is growing evidence against this notion. Rank-dependent utility theories relax this assumption by allowing decision weights to depend on the rank of an outcome among the set of all possible outcomes.

Luce (1991) and Luce & Fishburn (1991, 1995) propose and axiomatize a rank- and sign-dependent utility theory using an operation of joint receipt, or the simultaneous receipt of two or more objects. The utility of a risky or uncertain option is a weighted sum of the utilities of its component outcomes, where the weight of an outcome depends in a particular fashion on the rank order of the outcome and the sign of the outcome relative to the status quo. The utility function is assumed to be a negative exponential. Luce and his colleagues test the theory by studying individual axioms, and results have generally supported the theory (Cho et al 1994, Cho & Luce 1995).

In one test, Chung et al (1994) examine a property called event commutativity; the order of events should not matter to a decision maker as long as the outcomes arise under the same conditions (except for ordering). Violations would be problematic for the entire class of rank-dependent theories. Chung et al find solid evidence of event commutativity, consistent with both subjective expected utility and rank-dependent theories.

Tversky & Kahneman’s (1992) cumulative prospect theory is another rank- and sign-dependent representation that is identical to rank- and sign-dependent theory in all but two respects. First, it is based on a different axiomatization (Wakker & Tversky 1993), and second, it makes different assumptions about the utility function and the weighting function. In cumulative prospect theory, the utility function is a concave power function for gains, and a convex power function with a steeper slope for losses. The weighting function has an inverse-S form, first concave then convex.

These assumptions have not gone without criticism. Luce (1996) takes issue with the assumption of a power function for utilities; he shows that assumptions of cumulative prospect theory imply a negative exponential rather than a power function. The inverse-S shape of the weighting function has found support in some studies (Wu & Gonzales 1996, Tversky & Fox 1995, Fox et al 1996), but in a more general test, Birnbaum & McIntoch (1996) provide evidence against the inverse-S shaped form.

Lopes (1990, 1995, 1996) develops another rank-dependent theory called security-potential/aspiration (SP/A) theory. She argues that when subjects make risky choices, they are concerned with both security mindedness (avoiding the worst outcome) and potential mindedness (achieving the best outcome). Changes in attention to these goals influence the weighting function.
SP/A theory assumes that the weighting function is both “optimistic” for smaller probabilities and “pessimistic” for larger ones, a function she calls “cautiously hopeful.” The theory maximizes both a rank-dependent weighted average of the utilities and the probability of achieving an aspiration level.

Birnbaum and his colleagues have tested rank-dependent theories by examining branch independence: If two gambles have a common outcome for an event of known probability, the value of that common outcome should have no effect on the preference order induced by the other probability-outcome branches. This property must hold under expected utility theory, but not necessarily under rank-dependent theories. Birnbaum & McIntosh (1996) find violations of branch independence. Furthermore, the pattern is the opposite of that implied by cumulative prospect theory.

To account for these results, Birnbaum and others advance configural weight theory (Birnbaum et al 1992, 1997; Birnbaum & Beeghley 1997; Birnbaum & McIntosh 1996; Birnbaum & Viera 1997). Decision weights can vary with the rank and sign of the outcome as in the other theories, but also with the value of the outcome, the number of outcomes, the spacing of the outcomes, and the decision maker’s point of view. Testability comes from the assumption that utilities are invariant across all of these factors. Configural weight theory does not assume cumulative weighting. Cumulative weighting implies consequence monotonicity—if people prefer $X$ to $Y$ in one gamble, that preference should be in the same direction in the context of another gamble (Birnbaum & McIntosh 1996). Systematic violations of this property have also been found (Birnbaum & Sutton 1992, Mellers et al 1992b).

Some evidence suggests that the entire class of generic utility theories (Miyamoto 1988, 1992), including rank- and sign-dependent theories, does not capture preferences for gambles. Chechile & Cooke (1996) presented people with a reference gamble and two outcomes of a comparison gamble. Subjects were asked to adjust the probability of winning in the comparison gamble until the overall worth of the two gambles was perceived to be identical. In this task, special cases of generic utility theory could be represented as linear functions with an invariant slope across all reference gambles and a changing intercept. However, both the slope and the intercept systematically varied. The overall worth of a gamble depends not only on the properties that gamble, but also on the comparison gamble. Mellers & Biagini (1994) also show that the utility of an option systematically depends on the other option with which it is compared. Rank- and sign-dependent theories are insensitive to this dependency.

In summary, there is widespread agreement that risky and uncertain choices are rank dependent. But specific issues about the utility function and the weighting function—including shape, form (cumulative vs noncumulative), and factors that influence it—are still points of controversy.
Dynamic Decision Processes

Busemeyer & Townsend (1993) propose decision field theory to capture the deliberation process that occurs with conflicting values. The theory predicts the feeling of pleasure we may have about an important decision when the action is far away, and the later dread we feel when the action is imminent. Decision field theory also predicts preference reversals as a function of time pressure, violations of stochastic dominance, and the inverse relationship between decision time and choice proportions. This theory of decision making is similar to other accounts of cognitive processing, such as Link’s (1992) theory of perceptual discrimination and Ratcliff’s (1978) theory of memory retrieval. Diederich (1997) provides a generalization to the multiattribute case.

Certainty Equivalents

The certainty equivalent of a gamble is the amount of money for which a decision maker is indifferent between receiving the money for sure or playing the gamble. With judgment-based certainty equivalents, subjects state the worth of a risky option, often as buying prices or selling prices. With choice-based certainty equivalents, the point of indifference is inferred from a series of choices between a gamble and sure things.

Certainty equivalents differ from choices in two ways. First, they can produce different preference orders. Second, they can violate consequence monotonicity (Birnbaum et al 1992, Mellers et al 1992b). For example, Birnbaum et al (1992) found that subjects assign higher prices to a gamble with a 95% chance of $96, otherwise $0, than to a gamble with a 95% chance of $96, otherwise $24. This result has also been found with simple choice-based certainty equivalents (Birnbaum 1992) but not with direct choices (Birnbaum & Sutton 1992) or more complex choice-based certainty equivalents, such as those based on the PEST procedure (von Winterfeldt et al 1997). Differences between measures are important enough that any general account of preference should not assume they are equivalent; attempts should be made to describe similarities and differences among measures.

CONTINGENT VALUATION METHODS In the past two decades, contingent valuation (CV) methods have been used to measure the value of goods for which markets do not exist. People are often asked to state the maximum amount they would be willing to pay (WTP or buying price) to maintain a resource or return a damaged resource to the status quo. If CV responses represent economic values, they should vary with relevant factors and remain constant in the face of irrelevant factors (Baron 1997). However, CV responses are typically insensitive to the economic factors, such as the quantity of the good...

Some researchers argue that CV methods measure attitudes (Kahneman & Ritov 1994, Schkade & Payne 1994), moral sentiments (Kahneman & Knetsch 1992), or inferences about missing information (Fischhoff & Furby 1988). Occasionally, respondents are unwilling to answer questions. Baron & Spranca (1997) propose that these responses are evidence of protected values, a notion resembling taboo tradeoffs, that should not occur in economic theory. In sum, CV measures do not appear to be good measures of economic values, and better methods are desperately needed given the size of the monetary stakes.

CONCLUSION

Research on judgment and decision making has identified important limitations in cognitive and, more recently, emotional processing. The message of decisions errors and biases has had widespread effects. Confronted with real-world violations of rational choice theory, many economists and other social scientists now recognize the need for behavioral assumptions in the marketplace.

Within psychology, the concept of rationality is being reexamined. Errors are always defined relative to a normative framework, and within rational choice theory, definitions of errors have often rested on faulty assumptions. Researchers are now more circumspect about labeling behavior as irrational. Psychologists are examining alternative frameworks, and some theorists are moving away from rational choice theory to explore ways in which people find efficient, adaptive, satisfying decisions. For example, March (1994) proposes the notion of rule following. Rules have a generic form and can reduce the effort of tradeoffs. Many if not most of our decisions are of this type.

In sum, events occurring inside and outside of psychology are raising the level of debate about rational choice theory. Some social scientists now recognize the need for behavioral assumptions within their framework. Furthermore, psychologists, who have provided most of the behavioral violations, are more careful about what they call irrational.

The future of judgment and decision making research within psychology lies in its ability to develop other frameworks that make connections with research on emotions (Landman 1993) and cognition (Busemeyer et al 1995, Dougherty et al 1997), as well as on social and institutional factors (Kagel & Roth 1995). Such connections will undoubtedly influence our views of rationality and adaptiveness and provide fertile ground for research in the years ahead.
ACKNOWLEDGMENTS

Preparation of this chapter was supported by NSF grants SBR-94-09818 and SBR-96-15993. We are extremely grateful for the comments of Jerry Busemeyer, Ward Edwards, Gerd Gigerenzer, Duncan Luce, Lisa Ordóñez, Philip Tetlock, and Elke Weber during the preparation of this manuscript.

POSTSCRIPT

A tragic event that occurred during the period of our review was the death of Amos Tversky. Amos shaped the field in profound ways and made the field of judgment and decision making a much more exciting enterprise than it otherwise would have been. Ward Edwards, one of Amos’s graduate advisors, wrote an obituary in the Decision Analysis Society newsletter. Ward speaks for us when he says, “Certainly the care and depth of thought that shines through every paper on which Amos Tversky’s name appears, regardless of order of authorship, more than explains his overwhelming impact. The broad outlines of contemporary cognitive science show his fingerprints everywhere. We who will do the missing will continue to work in an intellectual environment of which he was a major designer. This is our good fortune, and his lasting achievement.”


Literature Cited

Becker GM, McClintock CG. 1967. Value: be-
Dawes RM. 1997. Judgment, decision making,


Luce RD. 1996. When four distinct ways to measure utility are the same. J. Math. Psychol. 40:297–317


Mellers BA, Cooke ADJ. 1996. The role of task and context in preference measurement. *Psychol. Sci.* 7:76–82


Redelmeier DA, Kahneman D. 1996. Patients’ memories of painful medical treatments:
Tetlock PE. 1992. The impact of accountability-