INTRODUCTION
The research domain generally referred to by the term JDM is vast and ill-bounded. It is, however, reasonably easy to identify the core concerns and issues it covers, even if one is unsure of the remote boundaries. The field has generally been concerned with choices made after some degree of deliberation: Choosing to take a particular job is included; choosing to remove one’s hand from a hot burner is not. The deliberation involved includes some prediction or anticipation, of two distinct sorts: prediction of the possible consequences of alternative actions and prediction of one’s evaluative reactions to these consequences. What will or might happen if I do A or B? And will I like these outcomes, or not? Selection of an action is often preceded by significant inferential effort, as when medical diagnosis precedes selection of a treatment. Substantial creative effort may be invested in generating action alternatives.

The term judgment often is used, imprecisely, to refer to several distinct parts of this process. The physician might use the phrase “In my medical judgment…” as a preface to a statement of what disease she thinks the patient is suffering from (diagnostic inference); what future course she expects the disease to follow (prediction or prognosis); what treatment she is recommending (decision); or what tradeoffs among risks, side effects, and prospects the patient will prefer (preferential prediction). Other topics often included under the JDM rubric are problem solving (viewing the physician as trying to solve the puzzle of the patient’s symptoms); information search (ordering tests, conducting exploratory surgery); memory (recall of earlier cases or symptom patterns); and dynamic decision making (as when the physician makes multiple interventions over time as the patient responds or fails to respond to treatments). JDM and its terminology, in short, is not neatly defined.

Given this inclusive and open-ended definition of the field and its constituent topics, we make no claim of comprehensiveness for this chapter, nor for the relative emphasis among the topics we have included. Our general goal has been to provide the reader with an introduction to the central issues in JDM, but we have been highly selective as to topics and relative emphasis. We have treated lightly or left out altogether many topics conventionally included in JDM surveys, in part by conscious (if inevitably biased) assessment of interest and research potential, in part by simple oversight. Our biases are generally toward actual or potential application rather than toward theory building per se. We note methodological issues only where they seem special to, or especially serious for, JDM. Finally, we have allowed ourselves a little scope for speculation on where the field might develop next, less in the spirit of confident prediction than in the hopes that it will spur our imaginations and those of others.

In this age of rapid and convenient electronic literature searches, we saw little point in stuffing this chapter full of exemplary citations on each topic. Other useful sources include two collections of papers sponsored by the Judgment and Decision Making Society: Goldstein and...

NORMATIVE/PRESCRIPTIVE VERSUS BEHAVIORAL/DESCRIPTIVE THEMES IN JDM

Perhaps more than other areas of the human sciences JDM research includes elements of both description and prescription, of trying to discover what people actually do when they form judgments and make decisions, and advising them on how they might do these things better. The advice-giving theme can be traced to mathematicians of the 18th century French court, who offered advice on such matters as the fair price for gambles (Bernstein, 1996; Stigler, 1986). The roots of the descriptive theme are more widely scattered, but were well established by the time of two landmark review papers (Edwards, 1954, 1961), which substantially launched behavioral interest in decision making.

The two themes seem to be built into the subject matter. If one starts, for example, with an interest in how a doctor makes a particularly difficult diagnosis (e.g., Einhorn, 1974), one would probably investigate the types of diagnostic information the doctor collects, the way she puts it together into an overall judgment, her ability to reproduce the same judgment on repeated cases, and so on. But it would be hard not to ask the evaluative questions: How well is she doing? Are her diagnoses correct? How well could anyone, or a computer, do in making this diagnosis from this information? How might she be helped to do it better? Conversely, a decision analyst might be able to show that, given specified preferences and probability estimates, a manager would be well advised to make a given set of investments. This still leaves open the manager’s ability to state appropriate preferences and to assess required probabilities—and to generate enough faith in the entire analysis to be prepared to take action based on it. Thus, serious descriptive work on decisions often reaches important normative questions, while intendedly prescriptive studies rise or fall on the realism with which they represent the psychology of the decision maker.

This interplay of descriptive and prescriptive issues is a central source of interest to many JDM researchers. However, it has also led to what is seen by many as an undue interest in decision errors. A major research program of the 1970s and 1980s, associated with Kahneman and Tversky (see section on heuristics and biases later in this chapter), assumes that observed decision behavior is generated by a reasonably small number of cognitive rules of thumb or heuristics, mental shortcuts that generally produce reasonable (and quick) results. These heuristics were demonstrated by showing that people generate systematic errors in specific, carefully constructed situations. The errors were defined as a deviation between what a subject did and the conclusions derived from some optimal rule—for example, a subject’s probability estimate when given some information and the estimate that would be generated by Bayes’s theorem in the same situation. This investigation of errors took on something of a life of its own (Edwards and von Winterfeldt, 1986; Jungermann, 1983), ignoring the facts that (a) the errors existed only if the optimal rule was, in fact, appropriate and accepted, and (b) there was little effort to assess the generality of the errors.

None of this is to suggest that humans are immune to decision error. Most of us, drawing on scientific evidence and personal experience alike, are happy to accept any help that is offered in our important life decisions. It is not clear, however, how common serious decision errors actually are. How might one assess an overall decisional batting average for the typical human, other than citing casual evidence suggesting it is not close to either 0 or 1,000? Without an agreement on what constitutes decision error, and an overall estimate of its frequency, one cannot assess how serious the biases caused by heuristic use might be. We argue only that, when presented with a normative recommendation, it is always wise to ask if its assumptions are descriptively accurate; and, when presented with a descriptive analysis of some decision maker, it is always interesting to ask how well he or she is doing.

INFERENCE PROCESSES

The Lens Model

Brunswik (1952) illustrated his discussion of visual perception with a diagram that has come to be called the Lens
Model (Figure 19.1). He argued that our skill at estimating some physical quantity such as the weight or distance of an object is the result of our ability to combine various imperfect “cues” to the quantity being estimated. For example, cues for distance include image brightness and sharpness, binocular disparity, parallax, and so on. None of the cues is perfectly correlated with actual distance, but a skilled perceiver can make use of the multiplicity and redundancy of cues to achieve highly valid estimates. The “lens” terminology simply draws attention to the similarity between the process of cue generation and integration and the diverging rays of light from an object being brought into focus by a convex lens.

Hammond (1955) proposed that the same model might be used to represent judgment processes. For example, the variable of interest might be a job applicant’s ability at some task, as reflected in cues such as scores on some predictive tests, reports from previous employers, and claimed experience in similar jobs. The judge’s task would be to combine these imperfect cues into an overall judgment of the candidate’s ability and thus into a prediction of the candidate’s performance on the job.

The great value of the lens model is that it draws our attention simultaneously to the judge (represented on the right-hand side as combining cues onto a judgment) and to the environment (represented on the left-hand side as some underlying state of interest spinning off imperfect cues). Achieving good accuracy requires both that the cues be reasonably informative about the underlying variable, and that the judge use these cues in an effective way. In fact, the mathematical relationships among the cue validities and utilizations and overall achievement have been helpfully analyzed in the so-called Lens Model Equation (Tucker, 1964). The model also draws attention to one of Brunswik’s methodological precepts, the call for “representative design” (Brunswik, 1955). In essence, this requires that cue-sets presented to subjects retain the cue ranges and intercorrelations found in some specified environment. Specifically, representative design forbids use of factorial crossing of cue values, since this procedure destroys naturally occurring cue intercorrelations. This will disrupt the judge’s normal judgment “policy” and may, in the limit, produce cue sets the judge finds incredible. Consider, for example, the reaction of an employer to a set of applicant records in which there was no relationship among test scores, undergraduate grade-point average, and quality of references. At least some of these applicants would probably be rejected as erroneous or fraudulent.

Multiple-Cue Probability Learning (MCPL) Studies

In more-or-less complete violation of “representative design” precepts, a large body of research has emerged broadly addressing subjects’ abilities to learn to use probabilistic information. The general format is to present the subject with a (long) series of trials in each of which several cues are presented and the subject is asked to predict the value of some criterion variable to which the cues are related. After the subject makes an estimate, he or she is told the correct answer before proceeding to the next trial. Such a format lends itself to endless variations in task characteristics: number of cues presented, their validity, the functional form of their relationship to the underlying variable the subject is to estimate, the quality of feedback presented, whether or not the task is embedded in a meaningful verbal context, whether or not learning aids are provided, and so on.
The evidence from dozens of such studies is that, except for the simplest versions, these MCPL tasks are very hard to learn. Simple generally means one or two cues, strongly and linearly related to the criterion, under conditions of low feedback error. For example, Slovic (1974) used a task with one linear cue that correlated 0.80 with the criterion, and found subject estimates approaching maximum possible performance in the last of 100 trials. However, when the cue validity was −0.80, learning after 100 trials was less than half this level. Deane, Hammond, and Summers (1972), using a three-cue task, found reasonable learning after 150 trials when all three relationships were positive, but almost no learning when the relationships were U-shaped. Learning improves somewhat when the subjects are warned about possible nonlinearities (Earle, 1970). Two-cue interactions are learned only if helpful verbal cues are provided (Camerer, 1981). Even after reaching high levels of performance under low-error feedback, subjects’ performance rapidly declines when feedback error levels are increased (Connolly & Miklau-sich, 1978). In short, as Klayman (1988) suggests, learning from outcome feedback is “learning the hard way.”

In many real-world tasks, of course, feedback is probably much less helpful than the “outcome feedback” provided in these MCPL laboratory tasks. A human resource professional trying to learn how to predict a candidate’s performance on the job from application materials receives feedback only after significant delay (when the applicant has been hired and on the job for some time); under high error (supervisor ratings may introduce new sources of error); and, crucially, only for those applicants actually hired (see Einhorn, 1980, on the inferential problems facing waiters who believe that they can spot good tippers). Laboratory MCPL tasks show excruciatingly slow learning of simple tasks under relatively good outcome feedback. Real-world tasks are almost certainly more difficult, and real-world feedback almost certainly less helpful, than the laboratory conditions. It thus seems unlikely that outcome feedback is the key to learning real-world tasks of this sort and interest in laboratory MCPL studies seems to have largely subsided in recent years.

Policy Capturing

Policy capturing, also known as judgment analysis (Stew-art, 1988), is the process of developing a quantitative model of a specific person making a specific judgment. The general form of such a model is an equation, often first-order linear, relating the judgments, J, to a weighted sum of the information “cues,” xi. Hundreds of such studies have been conducted, dating at least to Wallace (1923), who modeled expert judges of corn. Hammond and Adelman (1976) studied judgments of handgun ammunition, Slovic (1969) studied stockbrokers, Phelps and Shanteau (1978) studied hog judges, and Doyle and Thomas (1995) studied audiologists. In addition, policy capturing has been commonly used for organizational applications, such as decisions concerning salary raises (Sherer, Schwab, & Heneman, 1987), alternative work arrangements (Powell & Mainiero, 1999), cross-cultural differences in nonmonetary compensation decisions (Zhou & Martocchio, 2001), and applicant ratings and recommended starting salaries (Hitt & Barr, 1989). Policy capturing is thus a very widely used procedure.

It is also fair to say that the technique has been widely abused, and that many of the findings are hard to assess or interpret. The basic approach is so simple and obvious that it is easy to overlook some important subtleties that vitiate the final conclusions. We shall sketch some of these points here; see Stewart (1988) and Brehmer and Brehmer (1988) for fuller discussion, and Karren and Barringer (2002) for a focus on organizational studies using this methodology.

Suppose one were interested in modeling the judgment process of a university department head selecting candidates for graduate school. The department head reads an applicant’s file, writes a merit score between 0 and 100 on the cover, and moves to another file. At a later stage, the files are rank-ordered and applicants are admitted in descending order of merit score until all the places are filled. How might one model the department head’s judgment process?

A first step is to establish what information she is collecting from each file: the cues. Simply asking her what cues she is using may be misleading: It is possible that she is biased toward (or against) women, minorities, left-handers, or Scrabble players and is either unaware of the fact or chooses not to admit it. Second, how does she code this information? What counts as a “strong” grade-point average or an “acceptable” letter of reference? Significant work may be needed to translate the department head’s inspection of the file into a set of scale scores representing the cues she discovers and scores in it. Stewart (1988) provides helpful practical advice on this process, and Brehmer and Brehmer (1988) discuss common failures. Doyle and Thomas (1995) report an exemplary procedure for identifying cues, in their case cues used by audiologists in assessing patients for hearing aids. Once cues and judgments have been identified and scored, estimation of a standard multiple linear regression model is
straightforward. Interpretation, however, may not be. In particular, the interpretation of the relative weights given to each cue is conceptually difficult (see Stevenson, Busemeyer, & Naylor, 1991).

One subtle (and, in our view, unsolved) problem in policy capturing is how to meet Brunswik’s goal of representative design. This goal plainly prohibits constructing simple orthogonal designs among the cues: Such independence destroys patterns of cue intercorrelations on which expert judges may rely. Cue ranges and intercorrelations should reflect those found in some relevant environment, such as the pool of applicants or patients with whom the expert regularly deals. A sample of recent actual cases would appear to meet this requirement, but even here complexities arise. If one wishes to compare expert predictions with actual performance, then only the subset of applicants hired or admitted is relevant—and this subset will have predictably truncated cue ranges and intercorrelations compared to the entire pool. Changes in pool parameters arising from changes in the employment rate, prescreening, self-selection into or out of the pool, or even of educational practices may all affect the modeled judgment. The underlying problem of what exactly defines the environment the sample of cases is intended to represent is a conceptually subtle and confusing one.

Given these methodological worries, some caution is needed in summaries of research findings. Common generalizations (Brehmer & Brehmer, 1988; Slovic & Lichtenstein, 1971) include:

- Judges generally use few cues, and their use of these cues is adequately modeled by simple first-order linear models.
- Judges describe themselves as using cues in complex, nonlinear, and interactive ways.
- Judges show modest test–retest reliabilities.
- Interjudge agreement is often moderate or low, even in areas of established expertise.

In light of the methodological shortcomings noted above, we propose that such broad generalizations be taken as working hypotheses for new applications, not as settled fact.

**Heuristics and Biases**

Edwards (1968) ran the following simple experiment. He showed subjects two book bags containing 100 poker chips. Bag A contained mainly red chips, Bag B mainly black. He randomly chose one of the bags, drew out a small sample of chips, and showed them to the subjects. He then asked the subjects for their estimate of how likely he was drawing from Bag A. He found that subjects, initially persuaded that the probabilities were 50–50 before seeing the sample, generally revised their estimates in the direction suggested by the sample (i.e., toward A, if the sample was mainly red chips) but not as far as would be required by Bayes’s theorem. Edwards labeled the phenomenon *conservatism*. It involves three elements: a well-structured probabilistic task (e.g., sampling from two known populations); a sensible normative model for how the task should be performed (Bayes’s theorem); and an observation that actual behavior is systematically biased with regard to this normative model.

The dominant paradigm for research on judgment under uncertainty through the 1970s and 1980s, the so-called heuristics and biases paradigm (Tversky & Kahneman, 1981), was founded on observations of systematic errors of this sort: probabilistic tasks in which human behavior deviated systematically from a normative rule. The paradigm was, however, more than a simple catalog of errors. Tversky and Kahneman argued that the observed errors were manifestations of cognitive rules of thumb or heuristics, which, though generally effective and low cost, can be misleading in certain unusual circumstances. Thus, for example, we might guess the relative popularity among our acquaintances of various hobbies by noting the ease or difficulty with which we could bring examples to mind (the availability heuristic). This might work pretty well for most hobbies, but would likely mislead us for embarrassing or illegal hobbies, whose practitioners might well take pains to conceal their interest, or for praiseworthy hobbies, about which people would be likely to boast. Similarly, dramatic causes of death are judged to be commoner than less dramatic ones (Slovic, Fischhoff & Lichtenstein, 1979), and easily found words as more likely than those more difficult to search for (Tversky & Kahneman, 1973). (We discuss examples of heuristics and biases in prediction more fully in the following section.)

Work in this paradigm has declined in recent years. First, whatever the theoretical intentions, much of it became an ever-growing catalog of errors, with modest or no theoretical underpinnings that might allow prediction of when a particular heuristic would be evoked or error displayed. Second there was growing doubt about the appropriateness of some of the normative models invoked to demonstrate that errors had been made. Third, it became clear that at least some of the claimed “errors” were actually the result of subjects working successfully on problems other than the one the experimenter intended.
(See Jungerman, 1983, and Gigerenzer, 1991, for extended critiques of the heuristics and biases approach.) Research interest in documenting our shortcomings seems to have declined. Increasingly researchers are exploring the actual mechanisms that account for our performance, including the sometimes excellent performance of experts in real settings. (See Goldstein & Hogarth, 1997, and Connolly et al., 2000, for recent samplings of the literature.)

**PREDICTION**

**Simple Prediction**

There is evidence that, in making predictions, we use a variety of the heuristics discussed earlier. We will discuss three such heuristics: anchoring and adjustment, availability, and representativeness.

Imagine that an organization wants to predict sales for the coming quarter. A common approach would be to start with current sales as an initial estimate (the “anchor”), and then make an adjustment to account for market trends, new incentives, and so on. While this anchor-and-adjust heuristic may provide a reasonable estimate, research indicates that two potential problems may arise. First, the anchor may not be appropriate: If a new motivation program is applied to only a subset of salespeople, then the average of this group’s sales should be used as an anchor, rather than the average of all salespeople. Second, adjustment from the anchor may not be sufficient: The predicted value may be too close to the anchor of average sales. Bolger and Harvey (1993) found that decision makers used an anchor-and-adjust strategy for predicting events over time (e.g., sales) and that their adjustments were insufficient. Epley and Gilovich (2006) suggest that the underlying mechanism is inadequate search effort, at least when the initial anchor is generated by the individual, and may be overcome by incentives and forewarnings.

Another method for making predictions uses the “availability” heuristic discussed earlier: The likelihood of an event is judged by how easily instances come to mind through either memory or imagination. A manager may predict how likely a particular employee is to be late for work based on recollections of past episodes. However, availability may lead to biased predictions when we selectively attend to information that is available (e.g., a vivid or recent event) rather than consider historical/statistical data systematically. For instance, people who had recently experienced an accident or a natural disaster estimated similar future events as more likely than those who had not experienced these events (Kunreuther et al., 1978). Similarly, managers conducting performance appraisals can produce biased evaluations (either positive or negative) when they rely on memory alone: Vivid episodes and events within 3 months prior to the evaluation are over-weighted relative to other information (Bazerman, 1998). However, recent research suggests that people may be able to discount the biasing effect of availability when the cause of the bias is obvious to them (Oppenheimer, 2004; Sjoberg & Engelberg, 2010).

A third heuristic used in prediction is representativeness, in which the likelihood of an event is judged by its similarity to a stereotype of similar events. The “gambler’s fallacy” (for example, expecting a run of heads to compensate for an observed run of tails) appears to rely on the (false) belief that small samples of random events should accuracy reflect or be similar to the properties of the underlying distribution. Thus, a manager might predict the success of an employee by how similar he is to other known successful employees. Again, while this is generally a good initial estimate, using the representativeness heuristic can lead to systematic biases. First, people have a tendency to make nonregressive predictions from unreliable predictors—that is, to expect extremely high (or low) outcomes when the predictor is extremely high (or low). For example, Tversky and Kahneman (1974) attempted to teach Israeli flight instructors that positive reinforcement promotes learning faster than negative reinforcement. The flight instructors objected, citing examples of poor performance following praise and improved performance after reprimands. The instructors were attributing fluctuations in performance to interventions alone and not recognizing the effect of chance elements. Those trainees who received praise had performed at a level above their average performance, while those who were reprimanded had performed below their average. Statistically, both groups should tend to perform closer to their average performance on subsequent flights. Thus, the flight instructors falsely concluded that praise hurts and reprimands help because they predicted, by representativeness, that performance should be similar to the previous episode rather than regressing their predictions of performance to the mean. A parallel fallacy arises when we predict that the best-performing salesperson this year will be the top performer next year. Similar biases have been observed in stock market investment where, despite repeated warnings, investors seem to expect a stock’s previous performance to be highly predictive of its future performance (Chen, Kim, Nofsinger, & Rui, 2007; Rabin, 2002).
Another bias that has been attributed to using the representativeness heuristic is the tendency to neglect base rates or the prior probabilities of outcomes (Kahneman & Tversky, 1973). Imagine that a company knows that a small percentage (say 1%) of its employees is using illegal drugs. The company conducts a random drug test in order to determine which employees are using drugs and are subject to termination. The test is relatively accurate, being correct 90% of the time; that is, the test will be incorrect only 10% of the time when either a drug user tests negative (“false negative”) or a nonuser tests positive (“false positive”). Should the company fire employees who test positive for drugs? Most would say yes, thinking that the probability of being a drug user given the positive test result should reflect the accuracy of the test (90%). In fact, it is very unlikely ($p = 8.3\%$) that a person testing positive in this story is a real drug user. Although the test is relatively accurate, there are so few real users that most of the positive tests will be false positives. If we ignore the low base-rate of drug use, we hugely overestimate the test’s accuracy in identifying actual users. A similar example in a legal context, the so-called cab problem, is discussed by Kahneman and Tversky (1973).

Overconfidence

There are other potential problems in making predictions. In some situations, our judgments are overconfident. Experiments demonstrating overconfidence often ask difficult almanac questions in which subjects either choose between two options (e.g., “Which river is longer, the Tigris or the Volga?”) or state a range of values within which they are 90% confident a true value lies (e.g., “How long is the Tigris river?”). Klayman, Soll, Gonzalez-Vallejo, and Barlas (1999) found general overconfidence for almanac questions, but much more overconfidence for subjective confidence intervals than for the two-choice questions (approximately 45% vs. 5%). They found significant differences between individuals, but overconfidence was stable across individuals answering questions from different domains (e.g., prices of shampoo and life expectancies in different countries). A person who was overconfident in one domain was likely to be overconfident in another. Barber and Odean (2001) found men to be more overconfident than women in investment decisions. Men, as a result, traded more often and made less money than did women. Simon and Houghton (2003) found similar overconfidence, and significant financial losses, in managers’ new-product decisions. Overconfidence has been found in many, though not all, contexts (Moore & Cain, 2007; Yates, 1990). There is evidence that it declines with experience (Keren, 1987), and with instructions to think of ways in which an estimate might be wrong (Fischhoff, 1982). Overconfidence and its control has obvious implications in such organizational contexts as hiring, estimating timelines and costs, and developing business strategies.

There are also problems with learning from experience to make better predictions. The hindsight bias (Fischhoff & Beyth, 1975) hinders us in learning from our mistakes. In retrospect, we believe that we knew all along what was going to happen, and are unable to fully recover the uncertainty we faced before the event. This impedes learning the real relationships between decisions and outcomes that are necessary for good predictions. Unfortunately, simply warning people of this bias does not help, though inducing them to think of reasons they may be wrong can reduce the effect (Fischhoff, 1977). (Marks & Arkes, 2010, describe a new procedure based on source confusion that may help debiasing.) In addition, we may not seek the necessary information to test our beliefs since we have a tendency to seek confirming evidence (also called the confirmation bias; Wason, 1960) rather than disconfirming evidence. (See the section on information search, information purchase.) Finally, the structure of the environment may not readily provide information to test relationships since some information is naturally hidden. For example, personnel selection is often based on human resource test scores whose correlations with future job performance may be low. This will be true even for valid predictors of performance. We hire only applicants with high scores, so the variance of test scores for those hired is low and any variation in job performance will likely be due to other factors (e.g., motivation, training, random elements). We generally do not observe the performance of those we do not hire—data essential to testing the validity of our predictions.

Idea Generation

Before an outcome’s likelihood can be assessed, it must first be identified as a possibility. There is good evidence that we do not routinely generate many of the possible outcomes that may flow from our actions (Gettys & Fisher, 1979), and numerous remedial techniques have been proposed. One popular approach, group brainstorming, was first proposed in a nonacademic book (Osborn, 1953) as a way to generate as many ideas as possible. The participants were encouraged to improve, combine, and “piggyback” off other ideas without criticism in order
to generate more ideas than working individually. While this approach is intuitively appealing, subsequent research (McGrath, 1984) has shown that compared to brainstorming groups, the same number of individuals working alone (called nominal groups) produce more ideas with the same level of quality. Diehl and Stroebe (1987) concluded that the main reason appears to be production blocking: Since only one group member can talk at a time, the other members may forget their ideas, construct counterarguments, and so on in the meantime.

In the 1980s, computerized technology was developed to aid group brainstorming and decision-making processes (fortunately ignoring the evidence discussed above!). One popular system consists of several networked computers with a common main screen that can be seen by all in the room (Connolly, 1997; Nunamaker, Dennis, Valacich, Vogel, & George, 1991). Group members type ideas on their computers, and interact by passing files between machines. All members can thus be productive simultaneously, while drawing stimulation from reading and adding to one another’s files. This form of interaction appears to overcome the problems of face-to-face (F2F) brainstorming. Electronic brainstorming (EBS) groups can outperform equivalent nominal groups (Valacich, Dennis, & Connolly, 1994), at least when the EBS groups are large (approximately eight or more). It is not entirely clear why large EBS groups enjoy this advantage in idea generation (Connolly, 1997). Anonymity provided by the EBS system increases the number of ideas produced (Connolly, Jessup, & Valacich, 1990) and the number of controversial ideas (Cooper, Gallupe, Pollard, & Cadsby, 1998), but may decrease satisfaction with the task (Connolly et al., 1990). Several recent meta-analyses continue to provide evidence for the superiority of EBS over FTF brainstorming in terms of idea generation (Dennis et al., 2001; Dennis & Williams, 2005; DeRosa, Smith, & Hantula, 2007).

Interestingly, businesses continue to use F2F group brainstorming even though the literature clearly shows that it is inferior to both nominal groups and EBS. One reason may be its strong intuitive appeal. Paulus, Dzindolet, Poletes, & Camacho (1993) found that subjects predicted future performance and perceived actual performance as better in F2F brainstorming groups than in nominal groups, when in fact performance was superior in the latter. Another reason for the popularity of F2F brainstorming is the lack of access to EBS equipment. There is also some evidence that the performance of F2F groups can be raised to that of nominal groups by using highly trained facilitators (Oxley, Dzindolet, & Paulus, 1996). Kavadias and Sommer (2009) argue that the relative effectiveness of nominal and interacting groups is a function of both how structured the problem is and how diverse the group’s skills are. However, it may be that what researchers study (i.e., quantity and quality of idea generation) is not what business managers want (i.e., group well-being and member support). Dennis and Reinicke (2004) provide evidence that business managers use F2F brainstorming since it is superior in improving group well-being and member support and are less concerned with the increased idea generation capability of EBS.

**PREFERENCES**

**Values, Goals, and Objectives**

The idea of preference is fundamental to the idea of purposive choice: We prefer some possible outcomes to others and try to select actions accordingly. This is not the same as the claim that people “have” values (or preferences, goals, purposes, desires, etc.), in the sense that they can instantaneously say which of two real or imagined states they prefer at a given moment. As Fischhoff (1991) points out, some researchers (e.g., economists, opinion pollsters) behave as though people have fully articulated preferences for all possible objects and states of being, while others (e.g., decision analysts) suppose that we have only a few, basic values and must derive or construct preferences from these for most unfamiliar choices. An articulated values theorist might study a series of hiring decisions with a view to inferring the relative importance a particular human resource (HR) manager gives to different candidate attributes, such as experience, age, and gender. In the same context a basic values theorist might work with the manager to improve the accuracy or consistency with which her values are applied to future hiring decisions. (Indeed, it is possible to imagine doing both studies with the same manager, first capturing her “policy” from a series of earlier decisions, and then applying them routinely to subsequent decisions, a form of decision aiding called *bootstrapping*.)

Whichever view of valuing one assumes, there is plenty of evidence to indicate that the process can be imperfectly reliable and precise. Preferences for alternative medical treatments can shift substantially (for both patients and physicians) when the treatments are described in terms of their mortality rates rather than their survival rates (McNeil, Pauker, & Tversky, 1988). Subjects asked how much they would be prepared to pay to clean up one, several, or all the lakes in Ontario offered essentially the
same amount of money for all three prospects (Kahneman, Krench, & Thaler, 1986). Simonson (1990) found that people’s preferences for different snacks changed markedly from what they predicted a week ahead and what they chose at the time of consumption. Strack, Martin, and Schwartz (1988) found that students’ evaluation of their current life satisfaction was unrelated to a measure of their dating frequency when the two questions were asked in that order, but strongly related ($r = 0.66$) when the dating question was asked first. Apparently, the evaluation of one’s life overall is affected by the aspects one is primed to consider. MBA students’ ratings of their satisfaction with and the fairness of potential salary offers were markedly influenced by the offers received by other students in their class (Ordóñez, Connolly, & Coughlan, 2000). As these examples suggest, measures of preferences for real-life entities are sensitive to issues of framing, timing, order, and context and a host of other influences. It is unclear whether the problems are primarily those of imperfect measurement or of imperfect development of the respondent’s values and preferences themselves.

A common assumption of basic values researchers is that complex value structures are organized in the form of hierarchies or value trees (e.g., Edwards & Newman, 1982). The HR manager, for example, might consider a candidate’s attractiveness in terms of a few high-level goals, such as “job knowledge,” “motivation,” and “growth potential,” and assign some importance to each. At a lower level, these attributes would be decomposed so that “job knowledge” might include scores for formal education, job experience, and recent training, and so on. Such trees help to connect high-level values to lower level operational measures. More complex interconnections among value elements are also possible (see, for example, Keeney, 1992).

Utilities and Preferences

The term utility is used in two different ways in JDM. In the formal, mathematical sense (Coombs, Dawes, & Tversky, 1970), utilities are simply a set of real numbers that allow reconstruction or summary of a set of consistent choices. The rules for consistency are strict, but appear perfectly reasonable. For example, choices must be “transitive,” meaning that if you choose A over B and B over C, then you must also choose A over C. Situations in which thoughtful people wish to violate these rules are of continuing interest to researchers (Allais, 1953; Ellsberg, 1961; Tversky, 1969). Utilities, in this sense, are defined in reference to a set of choices, not to feelings such as pain and pleasure.

A very powerful formulation of this choice-based view of utility (von Neumann & Morgenstern, 1947) relies on the idea of “probabilistic in-betweenness.” Suppose A is (to you) the “best” in some choice set, and C is the “worst.” You like B somewhere in between. von Neumann and Morgenstern suggest that you would be prepared to trade B for a suitable gamble, in which you win (get A) with probability $p$ and lose (get C) with probability $(1–p)$. You could make the gamble very attractive by setting $p$ close to 1.0, or very unattractive by setting it close to 0.0, so, since you value B in between A and C, one of these gambles should be worth the same to you as B itself. The value of $p$ at which this happens is your “utility” for B, and expresses your preference for B in an unambiguous way.

The beauty of this approach is that it allows a decision maker to evaluate every outcome on a decision tree by the same metric: an equivalent (to her) best/worst gamble. Further, if some of these outcomes are uncertain, their utility can be discounted by the probability of getting them—their “expected utility.” If I value some outcome at 0.7 (i.e., as attractive to me as a best/worst gamble with 0.7 to win, 0.3 to lose), then I’d value a toss-up at that same outcome at $(0.5 \times 0.7)$ or 0.35. This provides a tight logic for expected utility as a guide to complex choices.

It is not clear how closely this formal view of utility conforms with the experience or anticipation of pleasure, desire, attractiveness, or other psychological reactions commonly thought of as reflecting utility or disutility. Indeed, the introduction of a gambling procedure for measurement gives many people problems, since it seems to involve elements of risk as well as outcome preferences. Many people turn down bets such as (0.5 to win $10, 0.5 to lose $5), despite their positive expected value (EV) $(0.5 \times 10) + (0.5 \times -5) = 2.50$, in the example). Why? One possibility is declining marginal utility: The $10 gain offers only a modest good feeling, while the $5 loss threatens a large negative feeling, so the 50–50 chance between the two is overall negative. This is referred to as risk aversion, though it may have little connection to the actual churn of feeling the gambler experiences while the coin is in the air.

The psychology of risk—what is seen as risky, how risk is talked about, how people feel about and react to risk—is a vast topic, beyond the scope of this brief chapter. Many studies (e.g., Fischhoff, Lichtenstein, Slovic, Derby, & Keeney, 1981; Peters & Slovic, 1996) raise doubts about our ability to assess different risks, and
show very large inconsistencies in our willingness to pay to alleviate them (Zeckhauser & Viscusi, 1990). Public policies toward risk are hampered by large discrepancies between expert and lay judgments of the risks involved (Fischhoff, Bostrom, & Quadrel, 1993; Slovic, 1987, 1993). The notion of risk aversion or risk tolerance as a stable personality characteristic guiding behavior across a range of situations finds little empirical support (Lopes, 1987). Recent brain-imaging work (e.g., Tom, Fox, Trepel, & Poldrack, 2007) has started to probe the neurological processes underlying risk taking and loss aversion.

Comparison Processes

The ideas we have reviewed so far all associate preference or value with an outcome in isolation from others. That is, they suppose that a specific outcome has a specific utility to a specific decision maker. Both casual reflection and careful research show that this assumption is false. One’s feelings about a $3,000 pay raise, for example, might shift significantly if one discovered a rival had made more, or less; if one expected nothing, or $5,000; or if it was given for merit rather than as a cost-of-living adjustment. Comparison processes of various sorts influence the value we attach to options and outcomes.

Inter-outcome comparisons are central to recent theories of regret and disappointment, which we will review below. Comparisons are also central to equity theory (Adams, 1965; Walster, Berscheid, & Walster, 1973), in which an outcome’s value is modified by the recipient’s judgment of whether or not it was fair. According to equity theory, equity is achieved when the ratio of outputs (e.g., salary, benefits, rewards, punishment) to inputs (e.g., hours worked, effort, organizational citizenship behaviors [OCBs]) is the same for all individuals being considered. Thus, in order to determine if equity is achieved, a comparison other (e.g., a coworker) is required. Early studies investigated equity theory by placing subjects in an experimental work context in which they received payment for the amount of work completed. Subjects were informed about the pay given to other, similar workers. Research results have strongly supported equity theory predictions (Greenberg, 1982). Equity imbalance was restored in a manner consistent with equity theory: Underpaid workers decreased their performance (i.e., lowered their inputs), whereas overpaid workers increased their performance (increasing inputs). In an interesting field study (Greenberg, 1988), workers were temporarily reassigned to offices that were either of higher or lower status than their regular offices. Consistent with equity theory, those assigned to higher status offices increased their performance, whereas those in lower status offices decreased their performance.

Choice Rules

In almost every practical choice situation, each of the options being considered has a number of features, attributes, or dimensions that affect its worth to the decision maker. A job, for example, might be defined in terms of such dimensions as salary, location, interest of work, promotion possibilities, and so on. Researchers have proposed a number of alternative models to describe the process by which decision makers choose between such multiattribute alternatives.

Multiattribute Utility Theory (MAUT) models suppose that what people do (or, in the prescriptive use, should do) is to evaluate each attribute of each alternative, add the resulting utilities into an overall utility for each alternative, and choose the alternative with the highest total. This is referred to as a compensatory model, in the sense that an improvement on one attribute can compensate for or trade off against a loss on another. (We discuss decision-aiding procedures for making these tradeoffs in the following section.) Some authors (e.g., Edgell & Geisler, 1980) have proposed modifications of the basic MAUT models (called random utility models) to reflect the fact that subjects’ preferences are not always stable from one occasion to another.

Conjunctive models reflect preferences of the screening type, such as an army physical examination. A candidate with flat feet, for example, would be rejected regardless of how well he or she scores on other measures of physical fitness. These models are thus noncompensatory, in the sense that more on one attribute may not make up for less on another: Any low attribute value makes the entire alternative low value. An early conjunctive model, the satisficing rule, was proposed by Simon (1955). Simon argued that, in real settings, MAUT procedures make unrealistic demands on a decision maker’s time and attention. Instead, decision makers search for an alternative that is acceptable on all important dimensions, and stop their search with the first such alternative. Note that this again introduces an element of probabilism into the choice, in that the order in which alternatives are considered may determine which of several acceptable options is found first. (Simon also argued that aspiration levels may change as search proceeds, adding a second element of probabilism.)
Lexicographic (dictionary-like) models rely on sequential comparisons between alternatives. Options are compared first on the modest important attribute and, if they differ, the winning option is chosen. If they tie, the next most important attribute is considered, and so on, until a winner is found. Another version of this, called the Elimination by Aspects (EBA) model, selects an attribute (or “aspect”) at random and eliminates from consideration any option that fails to reach threshold on this attribute. The process continues until only one option remains, and it is then chosen. (Note that neither of these processes is compensatory: Overall attractive options may be eliminated by a loss on an early comparison.)

Additive difference models (Tversky, 1969) assume that the decision maker compares alternatives along one dimension at a time, storing the sum of the differences favoring one alternative over the other. Probabilistic versions of this rule have also been proposed, in which comparison terminates when one alternative reaches some threshold of cumulative advantage over the other.

A number of authors (Beach & Mitchell, 1978; Payne, Bettman & Johnson, 1993) have suggested that the combination rule a decision maker uses represents a trade-off between effort and accuracy. The fully compensatory MAUT rule allows the fullest consideration of all attributes and values, but requires extensive information-processing effort. Other rules are less effortful, but do not guarantee that the best option will be chosen.

DECIDING: SINGLE-CHOICE EVENTS

Subjective Expected Utility Theory

The previous section discussed determining preferences among riskless options. However, selecting among risky options in which outcomes occur with some probability is even more difficult. For example, a firm may have to select between a set of new products to develop, each with probabilities of profits and losses. One of the simplest ways of placing a value on a risky proposition is by calculating its expected value (EV), which is the sum of the expected outcomes multiplied by their associated probabilities.

Clearly, such a calculation would be an imperfect guide to decision making in any single case. It can be easily shown that our preferences for risky propositions are not always consistent with an EV model. For example, how much would you pay for a gamble in which you flip a coin until the first head appears (on the $n$th flip) and pays ($2^n$)? If you get two tails followed by a head, you would receive $2^3 = 8$. Most people offer less than $4 to play this game. However, this game actually has an infinite EV, and according to the EV model you should be willing to pay as much as you are able. (The EV for the game is $\sum p_i x_i = \Sigma (1/2)^n 2^n = (1/2)2 + (1/4)$ $4 + (1/8)8 + \cdots (1/\infty)\infty = \$1 + \$1 + \$1\ldots$.)

Daniel Bernoulli (1738/1954) used the previous example (known as the St. Petersburg Paradox) to infer that people do not value a prospect in terms of the objective value of the outcomes, but on their subjective values or utilities. This model also explains why you might prefer $50 for sure over a gamble with a 50% chance of winning $100 and a 50% chance of $0 (a gamble with an EV of $50). Thus, the model of value is changed from EV with purely objective values of outcome value $(x)$ to the expected utility, EU, with subjective values (utilities) of outcomes, $u(x)$. EU $= \sum p_i u(x_i)$.

Later, Savage (1954) went a step further and proposed subjective measures of probability, too [i.e., subjective probabilities $s(p)$], so that prospects were valued at their Subjective Expected Utility (SEU) [SEU $= \sum s(p_i) u(x_i)$]. This model provided a way of placing value not only on risky events with monetary outcomes but also on uncertain events based on the degree of belief that an event will occur. SEU expanded the application of decision theory to include a much broader range of decisions.

Prospect Theory

Although EUT provides a good normative model of choice, several studies have demonstrated the theory’s weaknesses as a descriptive model of valuation and choice. The empirical violations of the axioms call into question the general applicability of EUT. For example, Tversky (1969) showed that, in certain problems, people consistently violate the transitivity axiom. The Allais Paradox (Allais, 1953) is a famous demonstration of how another EUT axiom (called Independence) is violated by many people.

Prospect Theory (Kahneman & Tversky, 1979) was developed to model how risky propositions are valued while accommodating decision behavior such as the Allais Paradox. The model uses the same general form 

$$EU = \sum s(p_i) u(x_i)$$
as EUT, but modifies the outcome value and probability functions to be more psychologically descriptive. A value of a prospect is defined as \( \Sigma v(x_i)\pi(p_i) \) where \( v(\cdot) \) and \( \pi(\cdot) \) are the value and decision weight functions, respectively.

The decision weight function, while similar to the subjective probability function of SEU, introduces new psychological features to subjective probability. One feature is that low probabilities are overweighted and high probabilities are underweighted. For example, Lichtenstein, Slovic, Fischhoff, Layman, and Combs (1978) have shown that people tend to judge low-probability health risks (e.g., botulism) higher than the objective values but tend to underestimate higher probability health risks (such as heart disease). Another feature of the decision weight function is that it is nonlinear. While objective probabilities sum linearly, decision weights do not. For example, \( \pi(.9) - \pi(.89) \neq \pi(.01) \).

In a second modification of EUT, Prospect Theory proposed a value function that was a significant departure from the previous utility functions (Figure 19.2). First, instead of defining subjective value with respect to overall wealth, the Prospect Theory value function defines value with respect to a reference point, often the status quo. Second, the value function for the domain of losses (below the reference point) is steeper than for gains. This leads to a result called loss aversion in which losses are more painful than equal magnitude gains are pleasurable. Finally, the value function is concave (risk averse) above the reference point, convex (risk seeking) below it. Since identical options can often be described in terms of different reference points, this raises the possibility that different ways of describing the same problem may shift choices from risk seeking to risk averse. This general framing problem is discussed in the following section.

**Framing**

To illustrate framing Hogarth (1987) presented MBA subjects with a choice between a riskless option, A, and a risky option, B with an EV equal to A. When the options were described in terms of money saved, A was preferred. When they were described in terms of losses, however, B was preferred. Due to the differing shapes of the value functions for the domains of gains and losses, people are risk averse when options are framed positively and risk seeking when options are framed negatively.

Another type of framing, attribute framing, has been shown for riskless options. For example, Levin and Gaeth (1988) showed that subjects evaluated ground beef more favorably when it was described as 75% lean than as 25% fat (though this advantage drastically diminished after consumers tasted the beef). The credit card lobby insisted on using the label “cash discount” rather than “credit card surcharge” when gas stations charged higher prices for customers using their credit cards instead of cash (Thaler, 1980). Levin, Schneider, and Gaeth (1998) provide a useful taxonomy of different framing effects. A study by Kuvaas and Selart (2004) suggests that such effects may result from negative framing stimulating more thorough and effortful information processing rather than simply changing the valence of the different outcomes.
Non-SEU Models of Decision Making

Most of the decision models discussed to this point have been variants on the expected value or expected utility model. They assume that a decision maker’s overall evaluation of some option is formed by an evaluation of the possible outcomes flowing from the option, discounting these evaluations to reflect the uncertainty of their occurrence, and then adding these discounted evaluations together to form the overall evaluation. From EV to Prospect Theory, the guiding spirit is evaluate, discount, and add. In this section, we look briefly at three models that do not follow this format.

Image Theory (Beach, 1990, 1993) sees the decision maker as concerned to maintain consistency among three mental images: a value image (summarizing her values and beliefs about rightness); a trajectory image (summarizing her goals and a path to their attainment); and a strategic image (a set of plans that guide tactical behavior toward the goals). The theory emphasizes screening of decision options for compatibility with the decision maker’s value image, and the selection of options to maintain consistency between the strategic and trajectory images. Actual comparative evaluation of options against one another (the “profitability test”) occurs only in the relatively rare case in which several options survive screening. Much of the research to date has focused on this screening process (Beach, 1998), with major emphasis on the number of “violations” an option must incur before it is rejected. There has been relatively little research on the nature and stability of the images themselves (Dunegan, 1993).

A second nontraditional decision model is presented by Lopes (1987, 1995) under the somewhat ungainly title of “security-potential/aspiration” (SP/A) model. The core intuition guiding the model is that assessment of an uncertain prospect such as a gamble generates a conflict between the downside or worst-case outcome(s) and the upside or best-case outcome(s). Some individuals (security minded) tend to be primarily concerned with the downside possibilities; others (potential minded) tend to be primarily concerned with the upside possibilities. For example, offered a choice between two gambles of equal EV, one with outcomes tightly clustered, the other with gambles widely distributed, the security-minded person will prefer the tight clustering (since the possibility of large losses is smaller) while the potential-minded person will prefer the wide distribution (since the possibility of large gains is larger). This basic balancing act is modified by the subject’s “aspiration level,” a level of gain or loss which the subject hopes to do better than. The model is consistent with a wide range of data, both from choices between gambles and from verbal protocols collected while making those choices. The SP/A model is a full alternative to Prospect Theory—and, indeed, does a better job of accommodating some of the evidence (Schneider, 1992).

A third non-SEU model has emerged from what is called the Naturalistic Decision Making (NDM) movement, which has been concerned with studying expert decision makers in their natural settings. These settings are often complex, time pressured, highly uncertain, high stakes, and dynamic, and thus unfriendly to thoughtful deliberative decisions (Orasanu & Connolly, 1993). Instead, researchers (Cannon-Bowers & Salas, 1998; Kline, 1993) have found that choice in such settings often turns on rapid assessment of the situation followed by rapid selection of an action that “matches” the situational demands. These recognition-primed or recognition-based (RPD) models (Cohen & Freeman, 1997) emphasize thinking much less, and rapid matching much more, than do conventional decision models. Indeed, these expert performances, though often highly effective, may address phenomena rather different from what has conventionally been called decision making. Experts doing what they know how to do may use mental processes quite different from those used by others struggling to find a course of action when they do not know what to do. Effective expert performance may not rely on reflective decision processes of the conventional sort.

Theory of Signal Detection

An important model of decision making that has been largely ignored in JDM research is the Theory of Signal Detection (TSD). Its roots are in efforts to guide early radar operators in deciding whether or not a given display included a “signal” (e.g., a real target) hidden in the “noise” on the radar screen. The TSD approach is driven by practical prescriptive goals of improving decision making, and is only indirectly concerned with the psychology of the decision maker. The approach is, however, of great generality for many applied problems, from assessing cracks in aircraft wings to detection of breast cancer, and from evaluating job candidates to testing for AIDS, drug use, or lying.

TSD (Getty, Pickett, D’Orsi, & Swets, 1988; Swets, 1988) considers a diagnostic situation, one in which repetitive choices must be made between two alternatives. An evidence system of some kind produces probabilistic
information of imperfect accuracy to guide the choices. For example, a test for some specific disease might produce a numerical score: If the score is high, the patient is likely to have the disease; if it is low, he or she is unlikely to have it. What should you, the physician, do with a given score? Since the test is imperfect, there is a possibility of an error either way. If you act as though the disease is present when it is not (a false positive), you incur costs of wasteful, painful, and perhaps dangerous treatments and patient distress. If you act as though the disease is absent when it is actually present (false negative), you incur costs of failing to treat real disease. You need to set a threshold on the test score at which you will act. The threshold requires consideration of how likely the disease is to start with (the base rate), and the costs and benefits of the two different sorts of error you might make.

The evidence system offers the decision maker a set of choices, which can be summarized in a plot of false-positive probabilities versus true-positive probabilities, called an ROC curve (Figure 19.3). The decision maker may decide to set a very strict threshold, insisting on a very high test score so that the chance of a false positive is small. The price she pays is that she will miss many true positives. Alternatively she could choose a lax threshold, acting even when test scores were quite low. Doing this would push the true-positive probability higher, but only at the cost of more false alarms. The ROC curve is thus a summary of the evidence system’s accuracy. A highly accurate system would offer very high true-positive probabilities with small false-positive probabilities. A completely useless system would offer identical probabilities of each. Anything that pushes the ROC up and to the left (higher true-positive probability for the same false-positive probabilities) represents an improvement in accuracy, and offers the decision maker a better range of options at which to set the threshold. Curve A thus offers a better menu of choices than does Curve B, and one research goal is to improve existing diagnostic systems in this way.

Independently of this improvement, it is possible to help the decision maker set appropriate thresholds, so as to make the best choice from those offered by the ROC curve. (Consider, for example, if you would want to use the same threshold on an HIV test for screening blood donations and for evaluating real patients. A false positive on the first case merely wastes a pint of good blood. In the second case, it would erroneously lead a patient to believe he had a life-threatening disease.) An excellent example of the TSD approach is Getty et al. (1988), in which the problem is improving the diagnosis of malignant breast cancers from mammograms.

Although it is easy to imagine the value of a TSD approach to a wide range of organizational decisions such as hiring, termination, new product development or R&D project selection, this potential does not appear to have been much tapped. [An exception is a report by Puranam, Powell, and Singh (2006) on due diligence procedures in corporate acquisitions.] Applications of TSD in organizations thus appear to be a promising research opportunity.

**EMOTION AND DECISION MAKING**

**Visceral Emotions**

Early research in decision making relied heavily on the “mind as a computer” model in which emotions were seen as irrelevant, or an active impediment, to rational choice. In the 1970’s and 1980’s Kahneman & Tversky’s (1979, 1984) more descriptive program showed how decision makers often use nonnormative heuristics rather than rational strategies, leading to systematic decision making biases. More recent research has started to explore the impact of emotions on decision-making processes.

Some of the earliest research in emotions and decision making investigated the impact of positive affect on cognition and information processing (Bower, 1981; Isen, Shalker, Clark, & Karp, 1978; Loewenstein and Lerner, 2002). Many of these studies simply categorized emotions as either negative or positive, without further differentiation. Negative emotions were found to lead to pessimistic expectations and to more analytical processing of information, positive emotions to more optimistic expectations and increased use of heuristics (Forgas, 2003; Johnson &
Tversky, 1983; Mayer & Hanson, 1995; Schwarz & Clore, 1983). More recently, Slovic et al. (2007) proposed the affect heuristic in which the positive and negative affective features of options shape evaluations rather than weighing of pros and cons.

Increasingly, research is showing that different positive or negative emotions have distinct effects on decision making (Ragnathan & Pham, 1999). Cognitive-appraisal theories of emotion (Lazarus, 1991; Smith & Ellsworth, 1985) differentiate between how emotions are experienced and their subsequent effects. Smith and Ellsworth (1985) proposed a six-dimensional taxonomy in which emotions are described in terms of certainty, pleasantness, attentional activity, control, anticipated effort, and responsibility. Associated with each emotion is a core meaning (or core appraisal theme) which summarizes the specific harms or benefits that arise in the environment. Lazarus and Cohen-Charash (2001), for example, suggest that the core theme of fear is facing an uncertain threat. These core themes increase the likelihood of specific courses of action (i.e., action tendencies). Thus, fear motivates the person to avoid potential harm (Smith & Lazarus, 1990).

The Appraisal-Tendency Framework (ATF) was proposed by Lerner and Keltner (2000, 2001) to connect cognitive-appraisal theories to judgment and decision making. ATF assumes that emotions trigger changes in cognition, physiology, and action. These changes generally help individuals respond to the event evoking the emotion, but they often persist beyond the eliciting situation. These emotion-related processes (also called appraisal tendencies) guide subsequent behavior and cognition in goal-directed ways, even in response to objects or events unrelated to the original cause of the emotion (Gasper & Clore, 1998; Lerner, Goldberg & Tetlock, 1998).

It is important here to distinguish between integral and incidental emotions. Integral emotions are directly connected to the decision task and may have a normative basis for affecting the decisions made. Thus, considering the regret you may feel about the outcome of various decision options may affect how you make your decision (e.g., Connolly & Reb, 2003). However, incidental emotions have no normative relevance to the decision task at hand since they are the result of outside events. There is no justifiable reason why academic achievements are given more weight in college admission decisions on cloudy days (when moods are negative) than on sunny days (when moods are positive) (Simonsohn, 2007).

ATF has stimulated a number of studies of the carry-over effects on decision behavior of situationally induced incidental emotions. Lerner and Keltner (2000, 2001) argued that two negative emotions, fear and anger, would have different impacts on risk assessment and risk-taking behavior. ATF proposes that fear involves appraisals of profound uncertainty: a sense that even basic needs are threatened by situational factors beyond one’s control. By contrast, anger involves appraisals of certainty and individual control: A demeaning offense occurred with certainty and the situation is under the control of human agency. Consistent with these predictions, Lerner and Keltner (2001) demonstrated that angry people perceived a given situation as less risky than fearful people did. In an earlier study, Ragnathan and Pham (1999) showed that anxious (fearful) individuals preferred low-risk/low-reward gambles but sad individuals were more likely to select high-risk/high-reward gambles. Kugler, Connolly, and Ordóñez (2012) confirm previous findings that show fearful individuals are more risk averse than angry people when the source of the risk is a chance event, but the effect reverses when the risk comes from the uncertain choices of others. Other research has shown that disgust can eliminate the endowment effect (Lerner, Small, & Lowenstein, 2004), guilt can lead to more cooperative behavior (DeSteno et al., 2010), and envy increases the use of deception in negotiations (Moran & Schweitzer, 2008). Thus, results indicate that distinct emotions of the same valence may have predictably different impacts on decision making.

Emotions and decision making is becoming a hot research topic in JDM. Peters et al. (2006) describe four different roles that affect can have in judgment and choice: information, common currency, spotlight, and motivation. Affect as information (Schwarz & Clore, 2003; Slovic et al., 2002; Loewenstein et al., 2001) proposes that we consult our feelings when making our judgments and choices, using feelings as valid inputs to the process. Taking this a step further, affective responses to options can serve as the “common currency” in which they can be compared. Models focusing on cognitive appraisals and action propensities (Lazarus, 1991; Lerner & Keltner, 2000; Zeelenberg et al., 2008) portray emotions as motivators for behavior. Finally, affect can act as a spotlight by focusing our attention to particular attributes of a decision problem.

Cognitive Emotions

Regret, disappointment, and related emotions are by far the most-studied of the emotions associated with decision making. We focus here on psychological research,
noting that economic choice theorists such as Bell (1982, 1985), Loomes and Sugden (1982, 1986), Irons and Hepburn (2007), and Bleichrodt, Cillo, and Diecidue (2010) have used the same terms but apparently referring to rather different concepts (see Connolly & Butler, 2006). In contrast to the economic approach, psychological research on regret and disappointment takes seriously subjects’ self-report measures of expected or actual emotional reactions to hypothetical scenarios (e.g., Connolly, Ordóñez & Coughlan, 1997; Gilovich & Medvec, 1994; Inman & Zeelenberg, 2002; Kahneman & Tversky, 1982) or actual events (Zeelenberg, Inman, & Pieters, 2001; Wroe, Turner, & Salkovskis, 2004).

In an early, much-cited study in this tradition Kahneman and Tversky (1982) gave subjects a brief scenario featuring two investors who each lose $1,200 as a result of owning a certain stock, Stock A. One investor initially owned a different stock, Stock B, but switched to Stock A. The other considered Stock B but decided to hold onto his Stock A. Subjects were asked to predict which of the two would experience more regret at his loss. An astonishing 95% of subjects felt that the investor who switched would experience more regret. This was interpreted as evidence that unfortunate outcomes resulting from action are more regretted than identical outcomes resulting from inaction. This so-called action effect was thought to lead to an “omission bias” in which, for example, parents would be deterred from vaccinating their child because they would judge a bad outcome resulting from vaccination (an action) as worse than a similar bad outcome resulting from not vaccinating (inaction) (Ritov & Baron, 1990, 1992; Baron & Ritov, 1994). (See Connolly & Reb, 2002, for a detailed critique of these studies.)

Subsequent research painted a more complex pattern. Gilovich and Medvec (1994) showed that action–regret linkages could reverse over time. Seta, McElroy, and Seta (2001) found the two-investors effect reversed if the protagonists were described as entrepreneurial businessmen rather than ordinary risk-averse savers. Zeelenberg et al. (2002) asked participants how much regret a soccer coach would feel if his team lost after he either changed or did not change his team. The active coach was seen as more regretful than the inactive coach, but only if the team had previously enjoyed a winning record. If they had been losing, loss after inaction was regretted more. Inman and Zeelenberg (2002) compared consumers who either switched brands or stayed with a previously purchased brand and were dissatisfied with their purchase. Predicted regret after switching was lower if the consumer’s prior experience with the initial brand had been poor, more regrettable than if prior experience with the initial brand had been good.

Drawing on these and other studies, Connolly and Zeelenberg (2002) proposed Decision Justification Theory (DJT). DJT proposes that decision-related regret has two components, one associated with an assessment of the outcome (“outcome regret”), the other with the decision process that led to the alternative chosen (“process” or “self-blame” regret). Outcome regret is driven by comparison of the actual outcome with some reference point (sometimes the outcome of an unchosen alternative, sometimes other reference points such as the status quo, the expected outcome, or the outcome received by another person: see, for example, Connolly, Ordóñez & Coughlan, 1997; Ordóñez & Connolly, 2000). Process regret, in contrast, is driven by the individual’s assessment of whether or not the decision was justified. For example, Zeelenberg et al.’s (2002) soccer coaches’ decision to change a losing team is seen as justified (and thus not blameworthy or regrettable). Changing a winning team is unjustified, and more regrettable. Seta et al.’s (2001) entrepreneurial investors were justified in taking action, and poor outcomes thus less regrettable, because that is what entrepreneurs do. Inman and Zeelenberg’s (2002) brand changers were justified in switching brands by their poor prior experience with the old brand. Reb and Connolly (2010) found that mothers whose vaccination decisions for their babies led to poor outcomes were expected to feel less regret when the decisions were based on a careful decision process. Indeed, although some justifications are specific to particular actions and roles (such as those of soccer coaches or entrepreneurs), the use of a careful, thoughtful, well-informed decision process—what Janis and Mann (1977) term a “vigilant” decision process—seems to be a regret-reducing justification across many contexts.

If poor outcomes are expected to be more regretted when they result from careless decisions, does the converse also hold? Does sensitizing people to possible regret motivate more careful decision processes? Recent evidence suggests that it does. Reb (2008) found that subjects primed to think about regret invested more effort, acquired more information, thought longer about their decisions, and made better final decisions than did those not so primed. Kugler, Connolly, and Kausel (2009) showed that regret priming can motivate more rational play in experimental games. Even quite subtle, unconscious priming of one or another type of regret can influence choice behavior. In a repeated decision task Reb and Connolly (2009) showed that unconscious priming of outcome regret led subjects to reject potentially painful feedback on the
outcomes of unchosen alternatives, impeding task learning and reducing final earnings (a trap they refer to as myopic regret avoidance). In contrast, subjects unconsciously primed for process regret accepted feedback (and thus the short-term pain of seeing that their outcomes could have been better), learned more, and performed better. Just as poor decision processes lead to increased regret, increasing the salience of possible regret can lead to improved decision processes.

Decision justification is central to Reason-Based Choice Effects (RBCEs; Shafir, Tversky & Simonson, 1993), such as the decoy effect (Huber, Payne & Puto, 1982), the accept/reject effect (Shafir, 1993), and the most-important-attribute effect (Slovic, 1975). In these RBCEs, “shallow but nice-sounding rationales” (Simonson, 1989, p. 170) can lead to nonnormative decisions. Recent work (Connolly, Reb & Kausel, 2010; Connolly & Reb, 2011) has shown that increasing regret salience can reduce or eliminate nonnormative decision RBCEs by prompting more scrupulous examination of one’s decision processes. Regret salience manipulations may thus constitute a relatively rare example of a theoretically grounded technique that effectively eliminates a class of decision biases and errors.

DECIDING: MULTIPLE RELATED EVENTS

Information Search, Information Purchase

One common way in which decisions are linked sequentially is when the outcomes of an earlier decision provide (part of) the information environment for the second. A doctor deciding on what laboratory tests to order for a patient is setting up the information environment in which she will make her subsequent diagnostic and treatment decisions. Similarly, a new product manager ordering a market survey is gathering information on which to base a later decision on whether or not to launch the product. In a shorter time frame, these acquisition and use processes merge.

Research on these processes has varied in how explicit the cost of acquiring information. Russo and Dosher (1983) recorded the subject’s eye movements to study which items of information he or she extracts from a decision table and in what order. The “cost” of an information item is the cognitive effort involved in attending to an item. A related methodology is the information board (Payne, 1976), in which decision-relevant information is displayed to the subject in a matrix of small envelopes that may be removed and opened. A computer-based analog called “Mouselab” has been extensively used (Payne et al., 1993) to explore underlying cognitive processes such as the combination rule being used by the subject.

Information cost is somewhat more explicit in work such as Wason (1960, 1968; Wason & Johnson-Laird, 1972) in which the subject makes an explicit request of the experimenter to turn over a card to decide whether or not an exemplar fits some unknown rule. In Wason and Johnson-Laird’s experiment, for example, subjects were shown four cards displaying E, K, 4, and 7. They were told that each card had a letter on one side and a number on the other and were asked which cards they would turn over to test the rule: “If a card has a vowel on one side, it has an even number on the other side.” Only 4% of their subjects selected E and 7, the correct choice. Almost half chose E and 4—an error, since the obverse of the 4 card cannot invalidate the rule, and thus produces, at best, evidence consistent with the rule but not testing it. This common finding has been interpreted as a general bias toward confirmatory search: seeking evidence that will confirm, rather than test, one’s initial beliefs. However, a penetrating analysis by Klayman and Ha (1987) suggests that such search patterns are better understood as examples of a “positive test” strategy, a generally appropriate heuristic that fails only in relatively rare situations, such as the four-card problem.

Explicit treatments of sampling cost flow easily from the Bayesian inference task discussed earlier (see the section on heuristics and biases). Instead of being presented with a sample of poker chips drawn from the selected bag, subjects are allowed to buy chips, at a fixed monetary cost per chip, before making their bet on which bag was selected, a bet for which they can win money. Findings from many such studies (see Einhorn & Hogarth, 1981, for a review) include:

- Partial sensitivity to normatively relevant variables. For example, Pitz (1968) found increased buying when cost per chip was reduced and diagnosticity was increased. Snapper and Peterson (1971) found some sensitivity to variations in information quality.
- Sensitivity to normatively irrelevant variables, such as information order (Fried & Peterson, 1969) and total information available (Levine, Samet, & Brahlek, 1975).
- Substantial losses (e.g., Kleiter & Wimmer, 1974), which persist with little or no learning over repeated trials (e.g., Wallsten, 1968).
- Both overpurchase and underpurchase (e.g., Hershman & Levine, 1970).
Largely parallel results are reported in an alternative, regression-based model of information purchase by Connolly and colleagues (see Connolly, 1988, for an overview).

The evidence from both Bayesian and regression models of information purchase suggest that subjects routinely and persistently make costly errors in balancing the costs and benefits of their information purchases. This should not be surprising. Optimal information purchase requires the subject to make accurate assessments of how accurate the different sources are, to select the best subset, and to combine the information acquired in an optimal way. Extensive evidence suggests that all three subtasks are quite difficult. It is thus likely that serious nonoptimalities will be found when the balance must be struck in practical settings. This is consistent with the reluctance of patients to seek second and third medical opinions before undertaking major courses of treatment, which, in our terms, represents a major underpurchase of decision-relevant information. It is also consistent with the huge body of evidence (Guion, 1975) on the predictive uselessness of unstructured job interviews—which are, nonetheless, still very widely used, and represent a huge overpurchase of decision-irrelevant information. Wherever information costs and benefits need to be brought into balance, then, there is good reason to suspect significant departures from optimality (March & Feldman, 1981). Applications range from improved design of Web sites (e.g. Peterson & Merino, 2003) to the impact of cell-phone use on drivers’ visual search patterns (Recarte & Nunes, 2003).

**Sunk Costs and Escalation of Commitment**

One important way in which a series of decisions over time can be linked is when nonrecoverable costs incurred at an earlier stage influence decisions at a later stage. The prescriptive advice on such matters is clear: The costs are “sunk,” and should play no part in the later decisions. Equally clearly, many of us violate such advice. We finish indifferent restaurant meals, sit to the end of bad movies, and remain in failed relationships so as not to “waste” the money spent on the restaurant bill or movie ticket or the time “invested” in the relationship. We fall, in short, into the “sunk-cost trap.”

Arkes and Blumer (1985) report 10 small experiments in which sunk-cost effects were demonstrated. Though most used a scenario format (and are thus open to the criticism that they involved the subjects in no real decisions), Experiment 2 made clever use of actual theater-ticket buying decisions to investigate sunk-cost effects. Of patrons buying season tickets for a local theater, one third paid full price, one third were given a modest discount, and one third a substantial discount, from the normal price. Patrons paying full price subsequently attended significantly more of the performances than did those who received discounts, though the effect faded later in the theater season. Arkes and Blumer interpret this as evidence that the larger sunk costs incurred by the full-price patrons influenced their later attendance decisions.

Similar effects have been reported in organizational (e.g., Staw & Ross, 1989) and other contexts (Brockner, Shaw, & Rubin, 1979). In a typical organizational study, Staw, Barsade, and Koput (1997) found that loan officers at banks were more likely to continue funding and extending problem loans when they had been responsible for the initial lending decision than when they took over responsibility for the loan after its initiation. A related effect in the persuasion literature, the “foot in the door” technique, involves winning compliance to a large request by first obtaining compliance to a smaller one (Freedman & Fraser, 1966). More subjects agreed to put up a large lawn sign when they had earlier been asked to sign a petition on the same subject than when subjects were approached directly with the large request.

Despite such apparently robust demonstrations, there is some confusion as to what phenomena are appropriately included in “sunk-cost effects,” and an embarrassing range of partially conflicting explanations has been offered. One setting in which escalating commitment has been demonstrated in scenario studies is in continuing to fund partially completed projects (e.g., Staw, 1976). However, when degree of project completion and expenditure are independently manipulated (Conlon & Garland, 1993; Garland, 1990), only the former factor shows an effect. Moon (2001) found that the two effects can operate independently of one another, and He and Mittal (2007) found that their relative impact changed over the course of a project. Public use of sunk-cost arguments by public officials may reflect either the entrapment of the speaker or the calculation that sunk-cost arguments will persuade the audience. Staw and Hoang (1995) claim to have demonstrated sunk-cost effects in their finding that basketball players drafted early (and expensively) into the NBA thereafter are played more and traded at higher prices than their performance appears to justify. The result could, however, simply reflect the failure of the authors’ performance model to capture just what a player is worth to a team. It is thus somewhat unclear just what is to be included as a sunk-cost effect, or how reliably such effects can be reproduced.
One account of the sunk-cost effect has been offered in terms of Prospect Theory’s loss function. The initial cost is taken as a loss (below the reference point), thus putting the decision maker into a region of risk seeking. Continuing the project now offers a risky project with some hope of gain, while abandonment forces acceptance of a certain loss (Thaler, 1980). Arkes (1996) and Arkes and Ayton (1999) argue instead for a quite general aversion to “waste,” a category mistakenly expanded to include partially completed projects or previously incurred costs. Staw (1976) and Aronson (1984) offer accounts based on self-justification, while Kiesler (1971) sees behavioral commitment as the central mechanism. Ku (2008) and Wong and Kwong (2007) tie escalation to decision-related regret, and Higgins (2002) offers an account based on his prevention/promotion framework. Brockner (1992) presents a multitheoretical perspective.

Overall, then, the sunk-cost effect and its relatives seem obviously worrying, possibly widespread, and open to a broad range of theoretical accounts. There is, however, a suggestion that we may be lumping together several rather different effects, each driven by a complex psychology of its own.

Dynamic Decision Making

Dynamic decision problems are those in which the decision maker may act repeatedly on an environment that responds to his or her actions and also changes independently over time, both endogenously and exogenously (Edwards, 1962). An example might be a senior manager’s efforts to improve low morale in an organization. She may, over a period of months, try a number of different interventions, scaling up successes and abandoning failures. Over the same period various factors internal and external to the organization may also affect morale. Clearly, such problems set decision makers extraordinary challenges.

They have also proved difficult for researchers, partly because of their inherent complexity, partly because of the experimenter’s partial lack of control. Complexity implies difficulty in deriving optimal strategies. Lack of control arises from the fact that the problem facing the decision maker at time $t$ is partially the consequence of his or her earlier decisions, as well as of the experimental conditions imposed. On the positive side, the growing availability of computers has helped both in the creation of realistically complex experimental environments and in the analysis of strategic alternatives. Some examples of the sorts of studies this allows include:

- **Simulated medical diagnosis:** Kleinmuntz and Kleinmuntz (1981) created a diagnostic task in which simulated doctors attempted to treat simulated patients on the basis of their initial symptoms and of the results of any tests the doctor chose to order. They could also act at any point to administer “treatments” which might or might not improve the patient’s health. Health fluctuated, over the 60 time periods of each trial, both in response to the doctor’s interventions and to the preset (downward) course of the disease. The simulated strategies explored included Bayesian revision, a heuristic hypothesis-testing strategy, and a simple trial-and-error approach. The computationally intensive Bayesian strategy yielded only modest improvements over the heuristic strategy in this environment, and even the simplistic trial-and-error approach did well on some cases. Further simulation results are reported in Kleinmuntz (1985), and experimental results with real subjects in Kleinmuntz and Thomas (1987).

- **Artificial worlds:** A number of European researchers (see Mahon, 2000, for a review) have explored dynamic decision problems with the aid of simulated worlds: firefighting in simulated forests (Brehmer, 1990), economic development in a simulated third-world country (Reither, 1981), control of a simulated smallpox epidemic (Hesse, 1982), and so on. Funke (1995) provides an extensive review, with studies classified as to the person, task, and systems factors each examined. Typical findings are those of Brehmer (1990) from his simulated firefighting task. Subjects initially perform quite poorly, but can learn this complex task with repeated play. Feedback delays impede learning substantially. Opportunities to offset feedback delay by decentralizing decision making were mainly ignored.

- **Systems dynamics:** A group strongly associated with MIT (Diehl & Sterman, 1993; Paich & Sterman, 1993; Sterman, 1987, 1989) base their dynamic decision making tasks on feedback dynamics models in which coupled feedback processes make response over time extremely nonintuitive to most subjects. For example, in Sterman (1987) subjects faced a capital budgeting task in which there was significant lag between ordering new equipment and having it available to meet increased demand. Most subjects in this task generated very large and costly oscillations, despite instruction in system linkages.

As this sampling suggests, empirical studies of dynamic decision tasks are difficult. The tasks themselves
are quite complex, even if greatly oversimplified versions of real-world analogs. Amateur subjects are thus easily overwhelmed, while expert subjects object to the unreality of the tasks. Findings thus tend to be task-specific and difficult to aggregate over different studies. Progress, clearly, is being made, but there are important challenges in this area.

MULTIPLE DECISION MAKERS

Organizations make many important decisions in groups, partly because the complexity of the issues requires multiple perspectives, partly because multiple areas of the organization want influence. (Organizational researchers commonly distinguish between groups and teams, but decision researchers refer to both as group decision making, a usage we will follow here). Group decisions may potentially improve decision quality, but at the cost of significantly more complex decision processes: Information must be shared, beliefs and preferences combined, and social interaction, conflict, and cooperation actively managed. In this section, we examine research addressing these issues for certain and uncertain outcomes, technology designed to aid group decision making, and negotiation between two parties.

Group Decision Making

Groups must communicate information if they are to improve decision quality. Thompson (2011) notes several possible impediments to information flow in teams: (a) message tuning in which the sender gives more or less information based on what she believes the receiver needs; (b) the sender lacks proper perspective taking and assumes the receivers know something that they do not (curse of knowledge), and (c) sender believes that other teammates know and understand their thoughts and attitudes (transparency illusion). Social factors also affect the group’s behavior. Senders may distort messages so as to be better received by the receiver, or use indirect speech (e.g., “The new VP of sales has an interesting strategy” vs. “I think the VP of sales is making some tactical errors”) to show deference to superiors.

Information flow is further degraded when a few team members do a disproportionate amount of the talking, known as the uneven communication problem (Shaw, 1981). Members may also display the common information effect (Gigone & Hastie, 1997), discussing only information they hold in common. In hidden profile tasks (Stasser, 1988), the best option can be overlooked unless members’ unique information is revealed. One remedial method is to require members to rank order the options rather than merely state their top choice (Hollinghead, 1996), which allows options with unique positive information to stay in the consideration set rather than being selected out early in the process. Hastie and Kameda (2005) tested nine different rules for combining member preferences over multiple options. Computationally intensive rank-ordering methods (such as Borda and Condorcet) performed very well, but simpler majority/plurality rules also performed surprisingly well. However, these rules are vulnerable to the hidden profile problem noted above.

Are groups better or worse decision makers than individuals? The answer depends on the situation and decision to be made (and, of course, on the criteria for “good.” In many settings a technically inferior decision to which the whole group as agreed may be an excellent choice.). There is no clear pattern of groups either reducing or increasing decision biases. Hindsight bias was slightly reduced with groups compared to individuals (Stahlberg, Eller, Maass, & Frey, 1995), though Bukszar and Connolly (1988) found no effect. Groups were even more affected than individuals by the representativeness heuristic in a base-rate (cab) problem (Argote, Seabright, & Dyer, 1986). And groups, like individuals, appear to be biased in their information search (Schulz-Hardt et al., 2000). Tindale (1993) argues that group effectiveness depends on the demonstrability of the problem. If one solution can be unambiguously demonstrated to be the correct answer, then the group will usually adopt it. Otherwise (as in the cab problem), the group decides by majority rule and individual errors are maintained (Tindale & Davis, 1985). Tindale (1993) presents data in which decision biases are reduced or enhanced by groups as compared to individuals. Groups may tend to be more overconfident than individuals (Fischhoff, Slovic, Lichtenstein, 1977). They may also be more economically rational, offering less than individuals in the ultimatum game and exiting more quickly in a centipede game (and, interestingly, earning less while doing so) (Bornstein, Kugler, & Ziegelmeier, 2004; Bornstein & Yaniv, 1998). Cooper and Kagel (2005) compared two-person teams to individuals in a signaling game and found that teams were more strategic, had higher outcomes, and transferred knowledge better in response to changes in payoffs.

Kerr, MacCoun, and Kramer (1996) reviewed studies of decision biases at both individual and group levels. They also conclude that decision biases can be either
smaller, equal to, or higher for groups as compared to individuals depending on the type of decision, the initial values of the individuals, and how individual values are aggregated into group decisions. They propose a formal model of group decision making, the Social Decision Scheme model (Davis, 1973; see the special issue of *Organizational Behavior and Human Decision Processes*, 1999, on this topic). This model links the aggregation rule for individual values and the decision rule used (e.g., “majority wins,” “truth wins,” or “all options equiprobable”) to the outcome selected. For example, Whyte and Sebenius (1997) found that groups did not debias individual estimates, which were improperly anchored on inappropriate anchors. Using symmetric differences squared (SDS) methodology, the authors showed that group estimates were based on the majority view that was biased before group discussion began. Finally, Yaniv (2011) showed that framing effects were eliminated in groups if the members were heterogeneous with respect to initial frame but were polarized if they all were exposed to the same frame before meeting as a group.

There are some conditions in which groups generally improve decision quality. Several studies indicate that heterogeneity (of attributes such as personalities, gender, attitudes, and experience) is positively related to creativity and decision effectiveness (Jackson, May, & Whitney, 1995). Guzzo and Waters (1982) found that the quality of group decisions and the number of diverse alternatives increased when expression of emotion was delayed until after alternative solutions were discussed. They suggest that early expression of emotions may reduce the group energy and narrow the range of accepted ideas. Under time pressure, quality of decisions generally declines, though task cohesion can help offset this effect (Zaccaro, Gualtieri, & Minionis, 1995). Finally, the popular book *The Wisdom of Crowds* (Surowiecki, 2004) provides anecdotal evidence that large numbers of people (crowdsourcing) can outpredict experts if individual opinions are diverse and independent; decentralized so that individuals can specialize and draw upon their local knowledge; and a method is provided for aggregating the individual judgments. However, Kostakos (2009) examined three popular voting Web sites including Amazon.com and found that the “crowd” typically includes a small group of experts that do the majority of the ratings. The wisdom of crowds may thus be heavily derived from the wisdom of a few experts.

Groups can also degrade decision performance. Janis (1972) coined the term *groupthink* to label “a mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members’ strivings for unanimity override their motivation to realistically appraise alternative courses of action.” A classic example is the failed Bay of Pigs invasion in which the American military sent Cuban exiles to overthrow the dictator Castro. These groupthink decisions are characterized by highly cohesive groups under high stress from an external threat and suffering low self-esteem from earlier failure or decision difficulty. Other attributes may also contribute: an illusion of invulnerability, collective rationalization, belief in the inherent morality of the group, insulation, lack of impartial leadership, direct pressure on dissenters, stereotypes of out-groups, and lack of established decision-making procedures. However, note that merely increasing group familiarity is not sufficient to cause groupthink: Watson, Michaelsen, and Sharp (1991) found that groups who spent more than 30 hours on decision-making tasks were more effective than individual decision makers.

**Group Decision Support Systems (GDSSs)**

Group decision support systems are designed to facilitate group decision making. GDSSs usually take the form of computerized, networked systems that aid in idea generation and decision making. A brief summary of key findings follows; a more detailed account can be found in Hollingshead and McGrath (1995).

In general, groups using GDSS demonstrate more equal participation and increased focus on the task than unaided groups but also interact less, take longer, have lower overall consensus, and report less satisfaction with the process and decision (Hollingshead & McGrath, 1995; McLeod, 1992). GDSSs provide a unique environment in which group members can interact anonymously. Jessup, Connolly, and Tansik (1990) showed that anonymous members using GDSSs tended to be more critical, more probing, and more likely to generate comments or ideas than when individual contributions were identified. For a very recent summary of past findings and future prospects for GDSSs, see Gray, Johansen, Nunamaker, Rodman, and Wagner (2011).

Do face-to-face (F2F) or GDSS groups make better decisions? The answer depends on the task. As indicated previously, GDSSs are better for idea generation. However, F2F interactions appear to be superior for problem solving and conflict resolution. Interestingly, Hollingshead and McGrath (1995) suggest that some of the benefits of GDSSs may stem from the structured aspects of the decision-making process rather than the GDSS itself. Shirani (2006) found that GDSS groups were more likely to
share unique information (i.e., to avoid the common information effect) than F2F groups. Archer (1990) found no differences in decision quality between GDSS and F2F when the decision process phases of a complex business situation were organized and managed in a rational manner.

Research on the behavioral impacts of GDSSs on group decision performance is still in the early stages and has largely used ad-hoc student teams. Research needs to be done on intact groups that have had experience working and making decisions together. In addition, as noted above, it may be that simply structuring the decision-making task can improve performance. However other features that GDSSs can provide may improve decision making in ways that cannot be achieved without them. A recent survey (Shim et al., 2002) indicates that organizations are increasingly making decisions in globally dispersed groups necessitating computer-mediated communication systems (CMCSs) and GDSS. This review indicates that F2F interaction is richer than CMCS and leads to many positive outcomes such as increased group cohesion, enhanced creativity and motivation, increased morale, fewer process losses, and better decisions. Given that organizations are increasingly using virtual teams that must interact entirely with CMCS, care must be taken to “foster interaction, inclusion and participation, which are all related to the feeling of ‘being there’ or social presence” (Shim et al., 2002). There is some evidence that virtual teams are less prone to escalation of commitment in a new product development context (Schmidt, Montoya-Weiss, & Massey, 2001), suggesting that there are contexts in which the lack of social richness can be valuable to group decision making. However, organizations should be cautious. Results of a meta-analysis (Baltes, Dickson, Sherman, Bauer, and Laganke, 2002) indicate that use of CMCSs decreases group effectiveness, increases time required to complete tasks, and decreases member satisfaction compared to F2F groups.

**Negotiation**

Negotiation is the process in which people determine “what each side shall give and take or perform and receive in a transaction between them” (Thompson, 1990). There is a vast literature in the field of negotiation and our review here is highly cursory. For further information on the psychological aspects of the negotiation process, see Thompson, Wang, and Gunia (2010), Bazerman, Curhan, Moore, and Valley (2000), and Tsay and Bazerman (2009). We will focus on dyadic negotiations; however, there is also an extensive literature in multiparty negotiations and coalition formations that we do not discuss here (see Crump & Susskind, 2008, and Murnighan, 1986, for reviews).

Early social psychological work in the 1960s and 1970s focused primarily on individual differences or situational characteristics. The extensive literature on individual differences has shown little effect on negotiations (Thompson, 1990). More recently, researchers have examined the interaction between individual differences and contextual variables. For example, Kray, Thompson, and Galinsky (2001) examine how men and women adopt different bargaining strategies after stereotypes about effective negotiators are activated. When stereotypes are activated implicitly, men are more assertive than women and men prevail in a distributive negotiation. However, women are more assertive (and more successful negotiators) than men when stereotypes are activated explicitly. In addition, other research (Babcock, Gelfand, Small, Stayn, 2006; Small, Gelfand, Babcock, & Gettman, 2007) indicates that women are less likely to initiate negotiations, but perform on par with male counterparts when they do.

The 1980s through 1990s used the behavioral decision research (BDR) as a framework. Raiffa (1982), in his decision analytic approach, shifted the attention away from prescriptions of optimal strategies to descriptions of actual negotiation behavior. Rather than propose optimal bargaining solutions based on objective facts of a negotiation, this type of research examines the perceptions of the situation, the other party, and the self. Thus, this format was not to present a normative picture of negotiations but to describe behavior and, at times, demonstrate the systematic deviations from the rational negotiator. In the 1990s, a social cognitive perspective was developed, with the focus on the negotiator as information processor (Thompson, Peterson, & Kray, 1995).

Many of the findings in this field have taken the heuristics and biases results (such as framing and overconfidence) and found them in a negotiation context. A great deal of evidence indicates that the framing of a negotiation has strong implications for negotiations. For example, in a labor–management salary negotiation (Bazerman, 1984), a raise from $10 to $11 an hour can be seen by labor as a gain of $1 or as a loss of $1 if the union demanded $12 an hour. Likewise, management can view $11/hr as a loss of $1, compared to the previous salary, or as a gain of $1, compared to the union’s demands. The greater impact of losses over equal-magnitude gains (i.e., “loss aversion”) results in a reluctance to trade concessions (Ross & Stillinger, 1991), creating a barrier to conflict
resolution. Neale and Bazerman (1985) showed that negotiators with positive frames were more likely to make concessions and were more successful than those with negative frames (however, negatively framed negotiators earned on average more per transaction when an agreement was reached). Real estate agents have been shown to anchor on the list price of a house and insufficiently adjust when assessing the value of a home (Northcraft & Neale, 1987), conflict management experts fall prey to the availability bias and do not search sufficiently for necessary information (Pinkley, Griffith, & Northcraft, 1995), and student negotiators were overconfident in believing their offer will be accepted in final arbitration (Bazerman & Neale, 1982).

Additional biases have been found that are unique to the negotiation context. One well-known bias, the fixed-pie assumption, occurs because the negotiators assume that they must distribute a fixed-pie (Bazerman, Magliozzi, Neale, 1985) rather than searching for integrated solutions that increase joint payoffs. This belief in the mythical fixed pie can also lead to the incompatibility bias (Thompson & Hastie, 1990, Thompson & Hrebec, 1996), in which negotiators falsely assume that their interests are incompatible with those of their opponents. Bazerman (1998) gives an example of a labor–management negotiation in which both sides value increased training programs, and thus, the workforce would be more flexible for management and lead to more job security for labor. However, due to the incompatibility bias, they settle for a less than optimal arrangement because they do not realize that they have common interests and negotiate as if a compromise must be reached. In addition, the fixed-pie assumption can lead to devaluing any concession made by the opponent (Ross & Stillinger, 1991): If management is offering more job training, it must not be too costly, or it must be benefiting them in some way.

Recent research augments the BDR perspective with a more cognitive focus (Thompson et al., 2010) that integrates subjective values of outcomes other than the negotiated agreement (Curhan, Elfenbein, & Xu, 2006) such as the relationship with the negotiating partner. For a recent review of the negotiation area see Thompson et al. (2010) for a general overview and Tsay and Bazerman (2009) for a decision-making perspective. One area that has received a great deal of attention recently is the impact of affect on negotiations (see Druckman & Olekalns in the 2008 special issue on emotions in negotiation). This can be further divided into emotions resulting from the negotiated outcomes (Galinsky, Seiden, Kim, & Medvec, 2002; O’Connor & Arnold, 2001), emotions spilling over from other events (Wood & Schweitzer, 2010), displayed emotions (Sinaceur & Tiedens, 2006), and the anticipated emotions of the opponent (Van Kleef, De Dreu, & Manstead, 2004).

CONCLUDING COMMENTS

As this selective survey of JDM connections to I-O psychology has, we hope, made clear, we see the linkage between the two fields as having accomplished significant work, but as having a potential for much more. As Highhouse (2002) points out, there are many topics in I-O that seem to fall naturally into the JDM domain: personnel selection and placement, job choice, performance assessment, feedback provision and acceptance, compensation, resource planning, strategic forecasting, and others. The two disciplines have, however, remained largely isolated, despite the clear potential for collaboration. Our hope is that the present chapter may contribute something to stimulating this linkage.

It may help a little if we clarify what we see as the current state of development of JDM. The mere name of the discipline makes an implicit claim: that there is sufficient commonality across different decision situations for a general theory of decisions to make some sense. We would assess the evidence to date on this point as mixed. Weather forecasters do have something to say to heart surgeons, and hog judges have something to say to HR practitioners; but it would be absurd to claim that we have a successful general theory of judgment and decision that embraces all four territories as mere applications. Any general claims require extensive local tinkering before they bring much insight to specific practical applications.

In our view the best contribution JDM can currently make to I-O issues is as a fertile source of interesting hypotheses, and as a provider of frameworks and instruments. For example, we would not read the literature on overconfidence in lab problems as supporting strong predictions that managers will be overconfident in predicting hiring needs. It does, we think, make such a hypothesis worth exploring. It also suggests how the relevant research could be conducted. In return, such research would inform JDM of the boundary conditions on its findings: When, for example, does overconfidence generalize, when is it bounded, what mechanisms are successful in minimizing it? It is this two-way enrichment of one another’s disciplines that we see as the potential for an enhanced collaboration between JDM and I-O. Our fond hope is that this chapter may do something to facilitate the interchange.
REFERENCES


