Chapter Ten

The Heuristics and Biases Approach to Judgment Under Uncertainty

Dale Griffin, Richard Gonzalez, and Carol Varey

Heuristics, Heuristics, and Heuristics

"Predictions are difficult to make, especially about the future." This statement, attributed by different sources to a United Nations official, Niels Bohr, and Yogi Berra, may be taken as self-excusing, self-mocking, or simply confused. Although most of us would agree that the world, both physical and social, is too complex to predict, we also have the experience of easily and effortlessly making many predictions. It is difficult to consider all relevant factors when evaluating the probability of a sports team winning, a stock increasing in value, or a relationship leading to marriage, but somehow when we consider such matters a feeling of certainty or uncertainty seems to "pop out" of the given situation. For example, on the day that we are writing this, a respected British politician was asked whether the current Kosovo peace talks would lead to a settlement; after a brief pause, he stated with confidence "the balance of probabilities are 40–60 against."

According to the "heuristics and biases" approach to human judgment, people typically use cognitive short-cuts that make probability assessments easy, but prone to error. Such short-cuts occur not only in predictions but in retrospective judgments of probability as well. Consider a recent article in a major British national newspaper. The article, titled the "20 million to 1 family," described how a couple had "broken all records by having eight children born in symmetrical girl–boy, girl–boy, girl–boy, girl–boy order." The heuristics explanation is that people incorrectly (but easily and effortlessly) judge that particular sequence to be extremely unlikely because the symmetrical pattern of births is extremely *unrepresentative* of a random series. Formal probability theory, in contrast, prescribes that any sequence of four boys and four girls is as likely as any other.

We thank Roger Buehler, Tom Gilovich, Peter Harris, Derek Koehler, Jason Riis, Norbert Schwarz, Winston Sieck, David Smith, and Frank Yates for helpful comments on a previous draft.

Based partly on their experience teaching statistics and on their observations of judgments and predictions in applied settings, Daniel Kahneman and Amos Tversky (Kahneman & Tversky, 1972; Tversky & Kahneman, 1973, 1974) proposed that intuitive judgments under uncertainty are typically controlled by judgmental "heuristics" rather than by the formal laws of probability. Kahneman and Tversky were not the first to suggest that classical "rational" models of statistical reasoning fail to describe actual human reasoning in many settings, but their program of research has been both more radical and more influential than most others. Their challenge to rational models influenced theory and research not only in cognitive psychology but also in social psychology, economics, political science, medical decision making, and legal studies. Some discussion of the three original heuristics, and a description of some of the classic example problems used by Kahneman and Tversky (summarized in Kahneman, Slovic, & Tversky, 1982), is a standard part of virtually all introductory textbooks in both social and cognitive psychology. Because of this ubiquity in the social psychological literature (see Sherman & Corty, 1984, for a review), we focus on the broader implications of the program and present a few selected examples in the Appendix.

There are at least three reasons why social psychologists should be interested in understanding the applications and implications of the heuristics and biases tradition. First, there is a fundamental tension in social psychology about whether to model human judgment as fundamentally rational or irrational (e.g. Asch, 1952; Nisbett & Ross, 1980), although the precise meaning of rationality is rarely defined. Second, models and explanations from the heuristics and biases program have been applied in social psychology to explain phenomena as diverse as causal attribution (Quattrone, 1982), self-perception (Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka, & Simons, 1991), egocentric biases (Ross & Sicoly, 1979), vividness effects (Moser, 1992), and risk perception (Sherman, Cialdini, Schwartzman, & Reynolds, 1985.). Third, many models currently in vogue in social psychology bear at least a superficial resemblance to the heuristics and biases approach, e.g. the cognitive miser metaphor (Fiske & Taylor, 1984); the heuristic–systematic model of persuasion (Chaiken, Liberman, & Eagly, 1989); the feelings-as-information model of Clore, Schwarz, and colleagues (e.g. Schwarz & Clore, 1988); and the stereotype-as-heuristic model (Bodenhausen, 1993), and a clarification of the overlap between models could be useful. The current review will address each of these three concerns. We first discuss the meaning of "rationality" that is most relevant to the heuristics and biases program, review the negative and positive messages of the original program, explore the chief criticisms of that program, and finally present a framework to organize the "second wave" of heuristics and biases research.

Heuristics in Historical Context

The rational model

The classical model of rational choice is central to the discipline of economics, and at its heart is the guiding principle of maximizing Subjective Expected Utility (SEU). According to this model, the "rational actor" assesses the attractiveness of a given option by evaluating the probability of achieving each possible outcome and combining that subjective probability with the subjective utility of each outcome. The rational economic actor then chooses the best option on the basis of the optimal combination of probability and utility. Economic theories that guide public policy in areas as diverse and important as taxation, environmental safety, and social security rely on the central assumption that individuals and organizations are rational in this sense. The behavioral work of Kahneman and Tversky (and many colleagues) questions the fundamental assumption of this model.

In a frequentistic definition, the laws of probability describe how to maximize the number of correct judgments over a large number of trials. The classical rational actor is expected to follow the basic rules of probability even for unique events, such as forecasting the probability of a recession in the year 2010 or the chance of peace talks succeeding. There are many events for which it is easy to calculate the "correct" probability (e.g. the chance of a given hand of cards). But in other cases, such as the prediction of peace in our time, the appropriateness of the probability judgment can only be tested by examining its *coherence* relative to other judgments (e.g. the probability of a subcategory must be smaller than or equal to its superordinate category) and by examining its *calibration* when aggregated together with several other judgments equated on probability (i.e. events predicted with .70 probability must occur 70 percent of the time). Note that coherence can be satisfied with regard to purely internal criteria, whereas calibration is specifically defined in regard to external criteria: how many things actually happened in the world. Violations of rationality in this model, then, do not imply anything about the relative importance of "hot" emotional versus "cool" cognitive factors, only about whether people follow the rules of subjective probability and evaluate their own preferences consistently.

To explain coherence, consider Bayes's Rule, which has been described as the "master rule" of categorical inference or categorical prediction (see Fischhoff & Beyth-Marom (1988) for a detailed and psychologically oriented discussion of Bayesian hypothesis-testing). Bayes's Rule defines how to use probability theory to update the probability for a hypothesis given some data. For example, when inferring the probability that a patient has heart disease (H1) on the basis of a positive diagnostic test (D), a rational physician would (implicitly or explicitly) calculate the following quantity, where H2 refers to the probability that the patient does not have heart disease.

$$\frac{P(H1|D)}{P(H2|D)} = \frac{P(D|H1)}{P(D|H2)} * \frac{P(H1)}{P(H2)}$$

The first quantity on the right-hand side is the likelihood ratio, which expresses the *relative* likelihood that a patient known to have heart disease would yield the test result *D* (for data) compared to a patient known not to have heart disease. The likelihood ratio thus expresses the *diagnosticity* of the given evidence *D*. In general, diagnosticity increases with increasing separability of the two competing hypotheses, increasing quality of the diagnostic data, and increasing sample size of the diagnostic data. For example, a given blood pressure reading would be more diagnostic in distinguishing between heart disease and a healthy heart than between heart disease and another vascular disease; it would be more diagnostic if it were taken by an experienced physician than by a beginning medical student; and it would be more diagnostic if it were based on the average of many readings than based on a single reading. The second quantity on the right-hand side is the odds ratio, or prior odds, which expresses the judge's belief about the relative prevalence of the two outcomes in the relevant population, that is, the relative probability of encountering a given member of each class (in the frequentist approach, the chance of encountering a given member from a random draw).

The strength of inference that can be drawn from a given set of evidence depends on the relative balance of the likelihood ratio and the prior odds ratio. If, for example, the diagnostic test has good validity such that the likelihood ratio is 9 : 1 in favor of heart disease, then a prior odds of 1 : 9 against heart disease leaves the rational physician with posterior odds of 1 : 1, or a .5 probability that the patient has heart disease. If, on the other hand, prior odds of 1 : 9 against are matched by a likelihood ratio of 1 : 9 against, then the posterior odds are 1 : 81 against, or a little over a .01 probability that the patient has heart disease.

Note that using Bayes's Rule to describe "ideal" probabilistic judgment in frequentistic settings, with repeated, exchangeable events such as drawing balls from an urn, is entirely uncontroversial. However, when Bayes's Rule is used to prescribe the updating of subjective probabilities about a unique event, some controversy entails (e.g. Savage, 1954). In particular, some statisticians argue that probability theory can only be applied to the frequentist case. However, as many applied researchers (including Keynes, 1921) have argued, if probabilistic statements about unique, real-world events are excluded from the domain of probability theory, nothing interesting is left. Wars, depressions, mergers, marriages, and divorces may happen with some regularity, but each is experienced as a unique event. Are probability judgments about such events without guidelines or standards? For now, it is enough that the classical economic model of rationality requires subjective probability judgments to follow Bayes's Rule.

Attacks on the rationality assumption

In the 1950s, inspired by the use of expert judgment in engineering systems developed during World War II, by the cognitive revolution that required human judgment to be modeled in terms of computer systems, and by the increasing contact between experimental psychology and economic decision making models, a number of research programs examined the issues of coherence and calibration in human probabilistic judgment. Herbert Simon (1957), early in his Nobel prize-winning research on economic models, argued that "full" rationality was an unrealistic assumption because of processing limitations in living systems (and, incidentally, in virtually all computers currently available). He proposed a limited form of rationality, termed "bounded rationality," that accepted the limited search and computational ability of human brains but nonetheless assumed that after a truncated search and after considering a limited subset of alternatives, people did act and reason rationally. For the moment, Subjective Expected Utility theory (and the underlying assumption of rationality) was safe, as long as it was modeled on a reduced set of stimuli.

Research by Ward Edwards (reviewed in Edwards, 1968) was designed to test the rationality assumptions more directly. From his research on how people revised, or "updated," their probabilities in the face of new evidence, Edwards concluded that people are not perfectly calibrated, but are generally coherent in their judgments. In particular, people do reason in accordance with the rules of probability (as summarized by Bayes's Rule) but they give new evidence too little weight and thus are "conservative." It is important to our later arguments to note that conservatism was only the most *common* finding in this research program. Systematic exceptions were found when participants were given new evidence of low probative weight; in this case, judgments were typically "radical," giving too much weight to the new evidence.

The work by Simon and by Edwards and colleagues is generally seen as the precursor of the heuristics and biases approach. However, there were several other flourishing research programs on subjective probability through the 1950s and 1960s, programs that cast further doubt on the rationality assumption. For example, Adams & Adams (1961) examined the calibration of subjects' probability judgments about their own knowledge, and found consistent "overconfidence:" for most probability levels, the actual percentage of correct answers was too low to justify the judged probability of success. Researchers using the Signal Detection model to study human perceptual judgments (e.g. Pollack & Decker, 1958) found that the correspondence between the rated probability of a "signal" being present and the actual probability of a signal depended on the difficulty of the recognition problem. When the stimulus was relatively difficult to recognize (e.g. because a tone was degraded with random noise or because a visual stimulus was very small), receivers' subjective probability judgments were too close to 1.0, that is, they were overconfident. When the stimulus was relatively easy to recognize, receivers' subjective probability judgments corresponded closely to the actual probability of receiving a signal and sometimes were even too low.

Throughout the 1950s, J. Cohen (e.g. Cohen & Hansel, 1956) studied intuitive conceptions of probability in children and adults, especially in terms of belief in "chance" and "luck" in gambling and risk-taking behavior. He concluded that intuitive conceptions of probability were qualitatively different than those described by the axioms of probability theory. Anomalies in conceptions of randomness noted by Cohen and others included two particularly robust phenomena: the gambler's fallacy and probability matching. The gambler's fallacy is the belief (implicit or explicit) that the "law of averages" requires that the probability of a given outcome of a chance device (e.g. tossing Tails on a coin) increases with a run of the alternate outcome (e.g. tossing Heads many times). Probability matching is the practice of predicting the more common event on a proportion of the trials corresponding to the base rate frequency of that event (e.g. if a roulette wheel was designed to end up "red" on 70 percent of spins, a probability-matching bettor would bet "red" on 70 percent of the trials, instead of betting "red" on each trial, which would maximize the probability of winning).

Also about the same time, Paul Meehl was describing two fundamental challenges to the optimality of clinical judgment. First, he noted that clinical prediction was almost entirely based on characteristics of the case being judged, with little or no concern for the relative prevalence or "base rates" of the possible outcomes (Meehl & Rosen, 1955). Second, he compiled a list of studies that compared the accuracy of clinical prediction with actuarial or formula-based prediction: formulas did better (Meehl, 1954). Some time afterwards, Oskamp (1965) demonstrated how trained clinical judges become increasingly miscalibrated (overconfident) as they gained more data about a case. Later, Mischel (1968) challenged the validity of clinical interviews to predict future behavior in very different situations. Most important for the present review, he pointed to the discrepancy between judges' beliefs and the empirical evidence of poor predictive validity.

These diverse findings and perspectives set the stage for Kahneman and Tversky's judgmental heuristics model of intuitive probability. The heuristics and biases program was not in any sense a deliberate attempt to account for the anomalies that littered the field of human judgment; it was simply an attempt to describe human judgment as experienced in the classroom and in the real world. Simon and Edwards had brought the potential conflict between normative rational models and descriptive human models into sharp focus, but had concluded that people were approximately or boundedly rational, within limits determined by their computational processes. However, there was considerable evidence that the assumption of calibration was generally untenable, and some evidence from Cohen's work that the axioms that predicted coherence were not consistent with intuitive judgments of probability. In this context, Kahneman and Tversky took a radical step: they proposed that the rules of probability, which define the rational "best guess" about outcomes, are not natural or intuitive methods of assessing degrees of belief or likelihood. Furthermore, they implied, simplifying the search set or restricting the number of computations was not enough to rescue the rationality assumptions. Instead, in many situations people naturally and spontaneously assess the likelihood of an outcome by processes that are qualitatively different from the rules of probability theory. In other words, "intuitive" judgment is not boundedly rational, but (at least in the classical sense) not rational at all.

Later critics have argued that the heuristics and biases program marked a sudden and arbitrary shift away from the past research on conservatism, which largely upheld the assumption of rationality (e.g. Lopes, 1991; Gigerenzer & Murray, 1987). This criticism is ill-founded, for, as we explain below, the heuristics and biases model is consistent with conservatism as well as with the other anomalies listed above. The heuristics and biases program accounted for the previous findings and also predicted many specific laboratory-based anomalies presented and tested in Kahneman and Tversky's early papers. We must emphasize that the laboratory-based demonstrations were never meant to be the phenomena to be explained – they were meant to illustrate the *processes* thought to underlie the phenomena of interest. The phenomena to be explained were judgments in the real world that seemed to be at odds with the dictates of probability theory.

Negative and Positive Aspects of the Heuristics Program: First Wave

From the first articles on heuristics and biases, Kahneman and Tversky noted that their program had two interrelated messages, one negative, about how intuitions do *not* work, and one positive, about how intuitions *do* work. In retrospect, it seems possible to identify two distinct stages of the program. In the first stage, the focus was on the surface structure of judgmental heuristics, and demonstrations were designed to show how case-specific information dominated intuitive judgment and led to the complete neglect of other normatively important information. The second stage (or as we describe it below, the "second wave") attempted to describe the deep psychological structure of judgmental heuristics, and the accompanying demonstrations were more likely to show how the (often conflicting) multiple sources of information were weighted.

A negative model of neglect

In the first stage, which dates from the original collaboration in 1969 to the 1974 summary paper, Kahneman and Tversky focused primarily on defining three judgmental heuristics (representativeness, availability, and anchoring-and-adjustment) by means of analogies with perceptual illusions. In simple, between-subject scenario experiments, Kahneman and Tversky demonstrated that people neglect prior odds ("base rates"), sample size, evidence quality, and diagnosticity, and instead rely on their immediate evaluation of the strength of the sample evidence to construct their subjective probability judgments. The experiments focused on everyday judgments and predictions about hospital births, school achievement, and professional membership, rather than abstract textbook probability questions about balls and urns, or dice and coins. Such a shift in context was neither irrelevant nor unplanned, as the authors noted that questions about chance devices were most likely to trigger the use of statistical rules rather than intuitive thinking. The authors acknowledged that almost any problem could be made "transparent" enough to allow participants to "see through" its purpose and therefore reason statistically, but argued that between-subject manipulations in non-chance settings were most informative about how people typically reasoned in everyday life.

The "negative" conclusion from this program of research – that intuitive judgments typically reflect only case-specific evidence, neglecting base rates and other features about the broader distribution – is enough to explain many of the anomalies in probability judgment listed earlier. If people focus only on the sample-specific evidence, then conservatism should be prevalent when base rates, sample sizes, evidence, and diagnosticity are high, and radical or overconfident judgments should prevail when they are low. This "psychology of evidential neglect" was implicit in the defining papers in the heuristics and biases program, and was later made explicit by Griffin and Tversky's (1992) "strength-weight" theory and then modeled by Brenner's (1995) random support theory. Koehler, Brenner, & Griffin (1999), using random support theory to model the neglect predictions, found substantial support for the basic neglect model in the everyday probabilistic judgments of physicians, economists, and lawyers working in real-world settings. Even weather forecasters, aided by computer projections and immediate outcome feedback, showed substantial neglect of base rate and validity considerations until they received specific feedback about their biases.

Criticisms of the "neglect" message began soon after the early laboratory studies were published. The initial focus of attention was the "lawyer–engineer" paradigm (Kahneman and Tversky, 1972), in which participants were given a personality description (see example 2 in the Appendix) and asked to judge the probability that the individual was an engineer rather than a lawyer. Although participants were also told the number of engineers and lawyers in the relevant population (70 vs. 30, or 30 vs. 70), the judgments reflected only the personality description; the base rates were neglected. Soon afterwards, one prominent critic claimed that he had "disproved the representativeness heuristic almost before it was published; and therewith . . . also disproved the base rate fallacy" (Anderson, 1996). In particular, Anderson had shown that base rates and case-specific information received about equal weight when manipulated across scenarios in a within-subject design. Kahneman and Tversky accepted that within-subject designs revealed the *capacity* for rule-based thinking whereas between-subject designs revealed the *actual* application of rules in practice. However, a later demonstration (Fishhoff & Beyth-Marom, 1984) cast doubt on the utility of within-subject designs to reveal much about reasoning capacity. When participants received a set of scenarios in which characteristics such as base rates were varied, such base rates were used whether or not they were normatively informative. It seemed that the participants were actively trying to make sense of the experimental game and felt that they should use what they were given, and especially that they should use what varied. Furthermore, when within-subject manipulations are examined more closely, it is found that participants combine the two types of information additively rather than multiplicatively, demonstrating that even when the base rates are made salient, they are not used in accord with Bayes's Rule (Kahneman, 1998; Novemsky & Kronzon, 1999).

Many economists, whose theories would suffer most if the heuristics challenge to classical rationality was widely accepted, wondered about whether the observed neglect biases would disappear with appropriate incentives or market conditions. In a series of studies, an economist (Grether, 1992) found that judgments consistent with the Bayesian model did increase very slightly, but significantly, with incentives for accuracy. More important, even in a chance set-up (balls sampled from bingo cages), with both sample evidence and base rates determined by drawing balls from a cage, there was still considerable evidence of heuristic thinking. The experimenter was puzzled to find that the direction of bias in the student economist subjects varied according to the size of the sample drawn and the discriminability of the two hypotheses under test. When the sample size was small and the two hypotheses were similar (so a sample of evidence had low diagnosticity), the data revealed apparent overconfidence which fit his definition of "representativeness effects." When the sample size was larger and the two hypotheses were very different (so a sample of evidence had high diagnosticity), the data revealed underconfidence or apparent "conservatism." However, as noted above, this pattern is consistent with the basic neglect model. Similarly, studies of business students in market games involving repeated plays and real incentives also revealed biased judgments in accord with the heuristic model (Camerer, 1987), but biases seemed to decline with repeated playing of the game.

Not to be outdone, social psychologists were quick to suggest a variety of ways to change the original scenarios demonstrating base rate neglect, from stressing the representativeness of the sampling procedure, to making the base rates more extreme, to making the target case less extreme, to changing the order in which the case and the base rate are received, to giving the case-specific information in a list rather than in an organized personality description, to varying the occupations of the people who put the personality descriptions together (see J. Koehler (1996) for a comprehensive list). Although none of these manipulations changed neglect into truly Bayesian judgment, they did demonstrate the familiar social psychological adage that people are active searchers for meaning (Griffin & L. Ross, 1991), and seemingly small changes in presentation and content can lead to marked changes in judgment (Tversky & Kahneman, 1982). It is important to note, however, that studies of social judgment in which people actively discovered the base rate for themselves (instead of deciding which of the experimenter's numbers was relevant to the task) also support a strong form of base rate neglect (e.g. Dawes, Mirels, Gold, & Donahue, 1993; Griffin & Buehler, 1999; Yates & Estin, 1999; Yates, Lee, Shinotsuka, Patalano, & Sieck, 1998).

The positive model: The original perceptual metaphor

Along with the negative message that people do not intuitively follow Bayes's Rule, Kahneman and Tversky developed a descriptive model of statistical intuitions. When people infer the likelihood of a hypothesis from evidence, they asserted, people intuitively compute a feeling of certainty based on a small number of basic operations that are fundamentally different from Bayes's Rule. In particular, these basic heuristic processes include computing the similarity between a sample case and the category prototype or generating mechanism (representativeness), computing how easily instances of the relevant category come to mind (availability), and adjusting an already existing impression or number to take into account additional factors (anchoring and adjustment). Thus, representativeness measures the fit between a case and a possible cause, or between a sample and a possible distribution. Availability measures the ease with which specific examples come into consciousness: a highly unlikely event is one that seems literally "unimaginable." Anchoring-and-adjustment is something quite different; it is not a measure, but a simplistic process of combination that fails to weight each component by its evidential value. These are heuristics because they are "short-cut" tools that bypass a more complicated and optimal algorithmic solution, where an algorithm is a step-by-step set of rules that guarantees a correct or optimal answer. Heuristics can be described in the language of "if–then" procedural rules. "*If* seeking the probability that a case is a member of a given category (or that a sample was generated by a given population), *then* compute the similarity between the case/sample and the category/population prototype." "*If* seeking the probability that an event will occur, *then* compute the ease with which examples of that event come to mind." "*If* a number is available for use and on the right scale, *then* adjust that number upwards or downwards according to knowledge that comes to mind." Whether such procedures were meant to be conscious strategies was not generally clear in the original papers.

Each of the operations described by Kahneman and Tversky yield a feeling or impression of certainty or uncertainty, but none of the heuristic operations are affected by some of the central features of the Bayesian algorithm, such as prior odds ratio, separability of the hypotheses, validity of the evidence, or sample size. Instead, these "direct assessments" of probability are fundamentally *non-extensional* and *non-statistical*, because they operate directly on the sample evidence without considering the relevant set-inclusion relations (the extensional rules), and without considering the degree of variability or uncertainty in sampling case information that is controlled by considerations of sample size and evidence quality (statistical rules).

In this approach, deviations from the normative model were not considered "failures of reasoning" but "cognitive illusions." This term emphasizes that the outputs of the judgmental heuristics, like the processes involved in vision and hearing, lead to compelling impressions that do not disappear even in the presence of relevant rule based knowledge. Furthermore, the heuristics do not represent a "strategy" chosen by the individual judge;

again like perceptual processes, the heuristics produced their output without guidance or active awareness of their constructive nature. This general notion was not novel; it had been introduced by J. Cohen (1960) in his study of "psychological probability:"

> Psychological probabilities which deviate from norms based on an abstract or "idealized" person are not errors, in a psychological sense, any more than optical "illusions" as such are errors. They can only be described as errors in terms of a non-psychological criterion. Knowledge of the objective lengths of the Muller–Lyler lines, for example, does not appreciably affect our subjective impressions of their magnitude. Precisely the same is true of the Monte Carlo fallacy [gambler's fallacy] . . . even mathematicians who are perfectly convinced of the independence of the outcomes of successive tosses of a coin are still inclined to predict a particular outcome just because it has not occurred for a relatively long time in a series of tosses. (Cohen, 1960, p. 29)

One of the novel aspects of the heuristics and biases approach was the deliberate strategy of creating "cognitive illusions" to demonstrate the heuristics at work; this naturally led to a focus on judgmental errors (as defined by the normative *non-psychological* model) in order to demonstrate the compelling nature of the heuristics.

The heuristic approach helped to explain existing anomalies in statistical intuition as well as predict new phenomena. In particular, the gambler's fallacy and probability matching can be seen as examples of representativeness at work. A long run is unrepresentative of a random chance process, and so we expect to see alternations to make the sequence seem more representative. In probability matching, the strategy of always predicting the most common outcome is completely unrepresentative of the kinds of patterns that seem likely to occur by chance, so predictions are made with the same kind of alternations that are representative of a random or chance process. Later, Gilovich, Vallone, & Tversky (1985) showed that people systematically misperceive random sequences because of the expectation that the sample sequence will "represent" the random nature of the causal distribution and contain many alternations and few long runs. When basketball fans were presented with a sequence of shots described as hits and misses, a majority perceived a sequence with a .5 probability of alternation as representing a "streak," because it included more long runs than they expected. An even larger majority perceived a sequence with a .8 probability of alternation as representing a "chance" sequence, because there were few long runs, and so the observed pattern matched the defining characteristics of a "random" process. Not surprisingly, such fans perceived actual players to be streak shooters, even though none of the players studied had shooting patterns that deviated from a simple independence model based on the assumption that hits were no more likely to follow a hit than to follow a miss.

The often-observed difficulties people have in understanding and identifying regression artifacts (e.g. Campbell & Kenny, 1999) also follow from the application of representativeness: people expect an effect to be just as extreme as its cause, regardless of the strength of the predictive relationship. Thus, children are expected to be just as tall, short, or clever as their parents, and experimental replications are expected to be just as significant as the original (significant) studies (see "Replicating a study" in the Appendix). Kahneman and Tversky (1973) coined the term "prediction by evaluation" to describe the process of matching the size of the effect with the size of the cause; the extremity of the causal variable is *evaluated* and then an outcome is *predicted* that is equally as extreme.

Furthermore, when, by the statistical law of regression that operates whenever predictive relationships are not perfect, children are less clever than their parents or replications yield weaker results than their originals, people invariably seek out causal explanations. Such findings have profound implications beyond the rejection of an unrealistic model of rationality. If people see random sequences as systematic deviations from chance, and develop causal explanations for phenomena that represent simple regression artifacts, we can expect an intellectual culture that develops and maintains unfounded superstitions and useless home medical treatments, that sustains multiple competing explanations of social phenomena, and distrusts the quantitatively guided conservatism of science.

Each of the three original heuristics has both process data and application data to support it (Heath, et al., 1994). Probably the most direct demonstration of representativeness is the Tom W. problem (see Appendix), in which participants were asked to predict the graduate concentration of an individual based on a sketch derived from a projective test. The predictions for Tom W.'s college major (measured in ranks) were negatively correlated with the base rate likelihood of each major listed, but were almost perfectly correlated with ratings of the similarity of the personality description to the prototypical college major. Note that in this study, no numbers were given by the experimenters to the participants or by the participants to the experimenters. The input variables were the personality description and the list of majors, and the output variables were three sets of rankings: base rate likelihood of majors, rated similarity between the description and the prototypes, and likelihood that Tom W. majored in each subject. More recently, Bar-Hillel & Neter (1993) showed that people not only rank the probability of category membership in order of the similarity of a description to its category prototype (i.e. by representativeness), but also are willing to bet according to representativeness, even when this violates the most basic rules of class-inclusion. Evidence consistent with prediction by representativeness has been observed in several applied domains, including foreign policy decisions and predictions of clinical psychologists and accountants (see Gilovich & Savitsky (1999) for a review).

The availability heuristic has been used to explain why people overestimate the probability of highly memorable risks, e.g. murder, and underestimate the probability of less memorable risks, e.g. suicide (Slovic, Fischhoff, & Lichtenstein, 1982). Although some critics questioned whether the ease of retrieving instances actually mediated the link between vividness and higher probability judgments (e.g. Shedler & Manis, 1986), Moser (1992) showed that ease of retrieval was a mediator, but needed to be measured appropriately using a self-generated measure of memory. In a series of studies, Schwarz and colleagues demonstrated that ease of retrieval – and not amount or content of retrieval – was, as Kahneman and Tversky surmized, the key determinant of availability effects (see Schwarz, Bless, et al., 1991).

The anchoring and adjustment heuristic also received close scrutiny in both laboratory environments (e.g. Cervone, 1989; Quattrone, 1984; Strack & Mussweiler, 1997; Wilson, Houston, Etling, & Brekke, 1996) and applied settings (e.g. Northcraft & Neale, 1987). These demonstrations confirmed that both explicitly random values and irrelevant values can serve as anchors and influence the final judgment of probability (as well as many other quantities). In one memorable applied demonstration, Northcraft & Neale (1987) suggested that even experienced real estate agents who should have ignored the asking price of a house anchored on this value when making evaluation estimates.

Critical perspectives on the heuristics and biases program

The prominence of the heuristics and biases program has made it a salient target for critics who object to its "negative" view of human rationality. The rather hot-blooded nature of the anti-heuristics backlash can be seen in the following quote, which is by no means the most extreme: "Are heuristics-and-biases experiments cases of cognitive misers' underachieving, or of their receiving a Bayesian hazing by statistical sophisticates?" (Barone, Maddux, & Snyder, 1997). Below we review and evaluate the major criticisms of the heuristic approach (see also Kahneman & Tversky, 1996).

Some of the criticisms have been aimed at the insular culture of the heuristics and biases tradition. These criticisms include: the approach is an abrupt and arbitrary departure from the boundedly rational conservatism model that preceded it (Gigerenzer & Murray, 1987; Lopes, 1991); the explanations are merely restatements of the phenomena (Gigerenzer & Murray, 1987; Gigerenzer, 1996); concerns about the generalizability of the heuristic and biases attack on human rationality (Lopes, 1991); and puzzlement over why only three heuristics were identified (Wallsten, 1980).

As described in the historical overview above, the heuristics and biases approach helps organize anomalies already observed, and can account for both conservative and overly extreme judgments. Ironically, given critiques about the atheoretical nature of the heuristics and biases approach, the perspective has more in common with modern conceptions of the mind such as connectionist models (Sloman, 1996; Smith, 1996) than it does with the modified humans-as-rule-based Bayesian models that came before it. Early citation analyses asserting that reports of poor judgment were over-cited in the literature (Christensen-Szalanski & Beach, 1984) have been overturned with more recent and comprehensive analyses indicating no "bias for bias" (e.g. Robins & Craik, 1993). The restriction to three basic heuristics made for a simple and elegant framework for thinking about judgment, which may have contributed to the immense impact of the original research paradigm. As understanding about judgment processes progressed, more heuristics have been added to the original list (e.g. causal simulation, Kahneman & Tversky, 1982a). We will return to the expanding scope of the program in our final section.

More recently, alternative frameworks have been proposed. Gigerenzer (1996) has criticized the lack of clear computational models underlying the heuristics and biases approach, and has developed his own set of judgmental heuristics based on computer simulations. These "optimal heuristics" (our term) are based on a satisficing model of judgment that goes back to Simon. Essentially, the two modern programs of heuristics differ in their guiding assumption: "fast and frugal" heuristics are based on the assumption that human judgment is optimal within processing limits, whereas heuristics and biases are based on the guiding assumption that unobservable heuristics should be no more optimal than the observed judgments used to explain them. Other frameworks, such as the problem-solving model of Ginossar & Trope (1987) and the adaptive decision making model of Payne, Bettman, & Johnson (1993), emphasize the flexible and goal-driven nature of judgment processes. Like dual process models of persuasion (e.g. Chaiken, Liberman, & Eagly, 1989; Petty & Cacioppo, 1986), the flexible or adaptive approach proposes that people strategically choose whether to save effort and use heuristic methods of judgment or to invest time and effort in using more complex, rule-guided processes. These dual process models differ from the heuristics and biases approach primarily in terms of emphasis. Even though the heuristics and biases approach explicitly acknowledges that people can represent and use abstract rules, its essential claim is that heuristic thinking is widespread even under the ideal conditions of high motivation, high ability, and high effort. It is possible to create a more inclusive model of heuristic thinking, as we attempt below (see also Fiedler, 1996; Dougherty, Gettys, & Ogden, 1999), but the message of the heuristics and biases program is that heuristic thinking is the standard and rule based thinking is the exception.

We next turn to two critiques that have attracted much research attention and raise questions about the fundamental underpinnings of heuristics and biases research. First, the claim that findings of the program are merely artifacts of the conversational rules between subject and experimenter, and second, the criticism that workers in this program have confused different definitions of probability. We take these two criticisms in turn.

Critics have claimed that many of the apparently "irrational" judgments observed in such studies were actually caused by rules of conversational implicature. There are two versions of this claim. The first is that people actively make sense of their environment, actively search for the appropriate meaning of questions, statements, conversations, and questionnaires, and that the same objective information can mean something different in different social or conversational contexts. This perspective is part of a *constructivist* approach to judgment (Griffin & Ross, 1991) that is consistent with the second wave of heuristics and biases research discussed below (e.g. Kahneman & Miller, 1986). Kahneman and Tversky (1982c) themselves discussed the problems with using what they called the "conversational paradigm" and noted that participants were actively involved in figuring out what the experimenters wanted to convey, just as if they were engaged in a face-to-face conversation. They further noted the relevance of Grice's maxims of communication to their problems (see chapter 9, this volume) and explicitly attempted to overcome the common-language ambiguity of terms such as "and" and "or" (Tversky & Kahneman, 1983).

However, this acknowledgment did not prevent a second and more critical version of the conversational perspective. The claim is that results of the scenario studies lacked external validity because changes in wording and context could reduce the rate of biased responses to questionnaire scenarios. For example, Macchi (1995) argued that base rate neglect may arise from textual ambiguity such that the verbal expression of P(D|H1) is interpreted as P(H1|D). Thus, the text "The percentage of deaths by suicide is three times higher among single individuals . . ." may be interpreted to mean "within the suicide group the percentage of single individuals who died by suicide is three times higher" (ibid., p. 198). To test this hypothesis, Macchi changed the key phrase to read "1 percent of married individuals and 3 percent of single individuals commit suicide" and found that this dramatically increased the number of participants who used both the base rate and the specific information. Of course, it is possible to re-apply a conversational analysis to the revised question, and it is difficult to know when the cycle should end. That is why it is so useful to have a real-world phenomenon to guide the evaluation of laboratory studies that otherwise can get lost in a perpetual cycle of "experiments about experiments."

The conversational perspective has also focused on the lawyer–engineer paradigm. Some follow-up studies challenged the explanation that the base rate neglect observed in the original paradigm was due to judgment by representativeness, and have been widely cited as evidence that heuristic thinking is eliminated in familiar, real-world social settings (e.g. Barone, Maddux, & Snyder, 1997). For example, Zukier & Pepitone (1984) found greater attention to base rate information when participants were instructed to think like scientists than when participants were instructed to understand the person's personality. A related study (Schwarz, Strack, Hilton, & Naderer, 1991) found greater attention to base rate information when participants were told that the personality sketch was written by a psychologist than when it was assembled by a computer. Thus, one might be tempted to conclude (despite the many other demonstrations of representativeness in the laboratory and the real world) that the use of statistical logic depends largely on social roles and contextual implications. However, a closer look at these studies leads to an interpretation more in line with a "constructive" perspective more congenial to the heuristics and biases approach. In both studies participants were presented only with a *low* base rate of engineers; inferences about base rate use were based on a paradigm that did not manipulate base rate. When, as part of our preparation for this chapter, we replicated each of these studies, crossing a base rate manipulation with the role/context manipulations, we found that within each base rate condition, the social context mattered, but across base rate conditions, the pattern of results was identical. In other words, even though people are sensitive to the role and context information, and use such information to shape their reaction to the personality description, they are still not sensitive to the statistical structure of the problem. Thus, these studies suggest not that judgment by representativeness is an artifact of a specifically contrived experimental situation, but rather that heuristics operate upon information that is actively constructed by the perceiver.

The second major critique is the claim that the heuristics and biases program is built solely on observations of probability judgments for unique events. Many defenders of the "objective" or "frequentist" school of probability have denied any role for the rules of probability in describing events that cannot be replicated for an infinite series. Nonetheless, it is undeniable that physicians, judges, and stockbrokers, along with virtually everyone else, use terms such as "probability" and "chance" to describe their beliefs about unique events. One of the greatest statisticians of the twentieth century has described the logical foundation of the subjective probability viewpoint as follows: "The formal rules normally used in probability calculations are also valid, as conditions of consistency for subjective probabilities. You must obey them, not because of any logical, empirical or metaphysical meaning of probability, but simply to avoid throwing money away" (De Finetti, 1970). We note that this point can also be made with respect to throwing lives away, or even throwing happiness away.

The frequentist critics of the heuristics and biases approach claim that when the classic demonstrations of heuristics are reframed in terms of aggregate frequency, the biases decline substantially or even disappear (e.g. Gigerenzer, 1994, 1998; Cosmides & Tooby, 1996; Jones, Jones, & Frisch, 1995). However, proponents of the heuristics and biases approach have explored this possibility for some time. For example, Kahneman and Tversky (1979) proposed that when making aggregate frequency judgments, people were more likely to recruit statistical rules of reasoning, especially rules of set-inclusion relationships, than when making individual probability judgments. Tversky and Kahneman (1983) proposed that set-inclusion relations were more compelling arguments when framed in frequentistic "counting" contexts. Griffin and Tversky (1992) proposed that aggregate frequency judgments led to greater attention to "background" information such as past performance (i.e. base rates). And Tversky and Koehler (1994) proposed that the violations of set-inclusion relations observed when compound hypotheses were explicitly "unpacked" into elementary hypotheses would be smaller for frequency than probability judgments. Thus, the dispute between critics and proponents of the heuristics and biases tradition arises not about whether probability and frequency judgments are psychologically distinct, or that frequency presentations are intrinsically simpler than probability interpretations, or even that the magnitude of biases are typically smaller in frequentistic formulations. The dispute is about the *causes* of the discrepancy and its implication for understanding the classic demonstrations of judgmental heuristics and heuristic thinking in real-world applications.

According to the heuristics and biases approach, the discrepancy between single-event probability and aggregate frequency judgments occurs because aggregate frequency judgments are less amenable to natural assessments that operate "holistically" on unique cases and are more sensitive to statistical or logical rules because the application of such rules is more transparent. Further, comparisons of the two tasks are difficult to interpret because different experimental artifacts affect each format (Griffin & Buehler, 1999). According to the frequentist advocates, a "frequency format" is consistent with the evolved software of the mind, and single-event "subjective" probability judgments are inherently unnatural (Gigerenzer, 1994, 1998). Supporting this perspective is evidence that people are extremely efficient, and seemingly unbiased, at encoding and storing the frequencies of letters and words they have been exposed to. On the other hand, this perspective cannot account for the observation that virtually all uses of the concept "chance" (meaning likelihood) in early English literature are consistent with a subjective, single-event judgment (Bellhouse & Franklin, 1997), nor that people untutored in Bayesian statistics regularly use expressions of subjective probability to describe their beliefs about the word.

The power of the frequentist critique rests on empirical demonstrations that judgmental biases "disappear" when aggregate frequency replaces single-event probability as the scale of judgment (e.g. May, 1986). Ironically, one of the first demonstrations of frequency effects on probability judgment was by Tversky and Kahneman (1983). They proposed that when a within-subject design was combined with a frequentistic presentation format, the conjunction rule of probability would be decisive over a heuristic answer. They created a conjunction scenario (the number of men over the age of 55 who have heart attacks) that could naturally be described in frequency terms: as predicted, when presented within-subjects the frequentistic version (but not the probability version) led participants to follow the conjunction rule (see "heart attack" in the Appendix). Gigerenzer (1991; see also Fiedler, 1988) replicated this finding, and concluded that frequency made the conjunction fallacy disappear. However, Kahneman and Tversky (1996) showed that conjunction effects consistent with judgment by representativeness were robust even with frequency judgments when manipulations occurred in between-subjects designs, reiterating that the combination of frequency and within-subject designs were necessary to create an "easy" or "transparent" version of the problem.

The frequency versus single event debate continues to generate new research questions, especially on the problems of base rate neglect and overconfidence (Brenner, Koehler, Liberman, & Tversky, 1996; Gigerenzer, 1994; Griffin & Buehler, 1999; Griffin & Varey, 1999). More importantly, the high volume of research activity, both by proponents and critics, suggests that the heuristic and biases approach is "healthy" and is still guiding research. These research problems have not "remained in the laboratory" but have branched out, with great influence, to applied areas (see Heath, et al. (1994) for a review of real-world applications). One reason why the area has remained fertile is because of a second wave of research that refocused the direction of the original work.

The Second Wave of Research: Heuristics Unbound

The original set of heuristics and biases demonstrations had a tremendous impact, both in terms of challenging existing theories and stimulating criticism. But as with any initial statement of a theory, there were some empirical anomalies left to be explained. Most prominent was the problem of "causal base rates" (Ajzen, 1977): when base rates could be given a causal interpretation (e.g. a high proportion of failures on an exam implied that a difficult exam *caused* the failure rate) they received substantial weight in judgment. This forced Tversky & Kahneman (1982) to include the computation of causality as a basic heuristic operation (see below), and to acknowledge that the distinction between case-specific and population based information was less sharp than originally proposed. This latter conclusion was reinforced by the finding that people were sometimes most responsive to the size of a sample *relative* to the size of a population (Bar-Hillel, 1979). Such a "matching" approach to sample size implied a broader kind of representativeness calculation, or as Bar-Hillel termed it, a second-order representativeness. The sharp distinction between heuristics that operated on cases, and rules that operated on abstract statistical quantities, it appeared, was not always so sharp, and seemed better captured by a more flexible distinction between "holistic" and "analytic" thinking. Furthermore, the initial statements of the heuristics and biases approach contained some ambiguity with regard to whether judgmental heuristics were deliberate strategies to avoid mental effort, or largely automatic processes that were uncontrolled and uncontrollable. These issues were addressed by a second generation of papers on judgmental heuristics by Kahneman and Tversky.

The second wave of heuristics research began with an analysis of the "planning fallacy," the tendency for people to make optimistic predictions even when aware that similar projects have run well over schedule (Kahneman & Tversky, 1979). This paper introduced a new form of the perceptual metaphor, contrasting an "inside" and an "outside" perspective on prediction. Using an inside or internal perspective, a judge focuses on the specific details of the current case; using an outside perspective, a judge "sees" the specific case as one instance of a broader set of instances. Shortly afterwards, a paper on causal reasoning (Tversky & Kahneman, 1982) demonstrated how intuitive or heuristic processes could be applied to both case-specific and distributional information as long as both types of information were in a form amenable to "natural assessments." For example, base rates that have causal implications (e.g. a sports team has won nine of its last ten games) may induce a computation of a "causal disposition" (Kahneman and Varey (1990); e.g. see "Blue cab/green cab" in the Appendix). These two approaches blur the sharp distinction between case-specific and statistical information and instead distinguish between information that can be directly evaluated by natural assessments in a holistic manner and information that requires logical inference before it can be used.

Two key papers in this second wave of research included the exploration of the "conjunction fallacy" (Tversky & Kahneman, 1983), which introduced the notion of low-level natural assessments, and the statement of "support theory" (Tversky & Koehler, 1994), which described how assessments of evidential support are translated into probability judgments. Although cited primarily for the memorable "Linda problem" (see Appendix), the 1983 paper further developed the perceptual model of judgmental heuristics and clarified the role of abstract rules in intuitive statistical judgment. In this and related papers (e.g. Kahneman & Tversky, 1982a; Kahneman & Miller, 1986), Kahneman, Tversky, and colleagues distinguished low-level "natural" or "routine assessments" that are relatively automatic and spontaneously evoked by the environment, from higher-level judgmental heuristics, which are typically evoked by an attempt to answer a question. The clearest candidates for natural assessments are computations of similarity, causal potency, and counterfactual surprise.

Tversky & Kahneman (1983) chose the conjunction rule of probability as a case study in the conflict between heuristic thinking and rule based reasoning. They argued that the conjunction rule of probability (no conjunction of events can be more probable than either constituent event alone) is the most basic and compelling rule of probability and is understood, in some form, by virtually every adult. Thus, in a wide variety of contexts, they examined when the conjunction rule would overcome the "conjunction fallacy," the tendency to judge a conjunction as more probable than its least likely constituent. They made conjunctions seem likely by using representativeness (the combination of events or descriptions were more similar to the target description than one or both of the constituents), availability (the combination of events or descriptions were better search cues than one or both of the constituents), and causal relatedness (the combination of events created a causal link that seemed plausible, easy to imagine, and therefore more likely than one or both of the constituent events). The real-world phenomenon that is reflected in the conjunction fallacy studies is that as predictive scenarios become more detailed, they become objectively more unlikely, yet "feel" more likely. The authors noted that many participants reported being simultaneously aware of the relevance of the conjunction rule and the feeling that the conjunction was more likely than the constituent categories. Conjunction fallacies were extremely common in between-subject designs, quite common in non-transparent within-subject designs, and only reduced by a combination of a within-subject design and a frequentistic design in which participants could see that the *number* of people with *A&B* must be less than the number of people with *A*. Except in special circumstances, then, heuristic thinking overwhelms the rules of probability, even when those rules are known and endorsed by the intuitive judges.

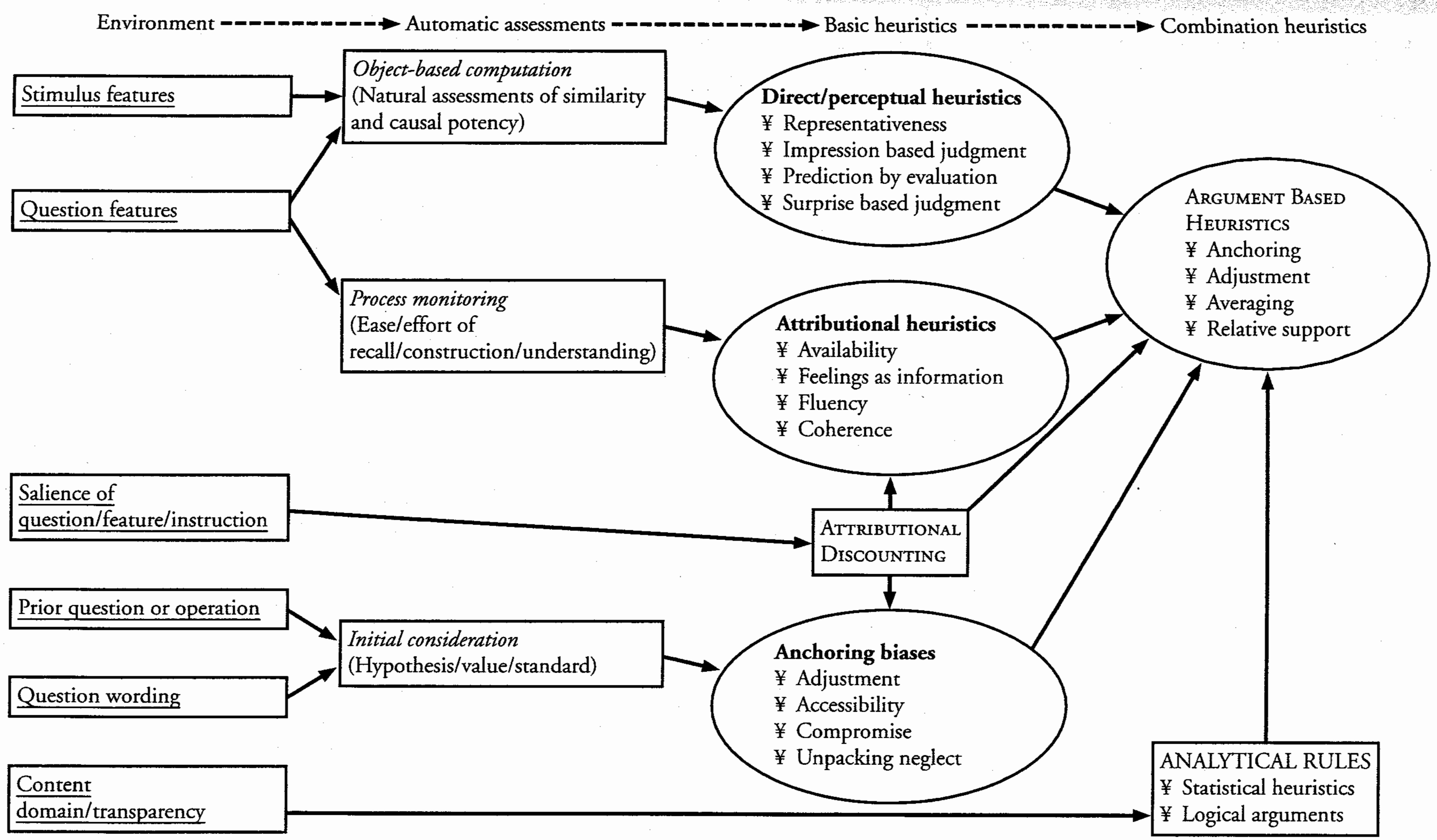
Note how this model is fundamentally different from the "cognitive miser" model that dominates current social cognition. Heuristic judgments are not explained as the result of too little thought due to cognitive laziness or reduced motivation, but as the result of "thinking too much" in quick and natural ways. This model of spendthrift automatic processes was termed "mental contamination" by Kahneman & Varey (1991), who related the basic processes of heuristic thinking to a wide range of perceptual, cognitive, and social examples, including the Stroop effect and motor effects on persuasion (Varey, 1991). Wilson and colleagues have also used the same term, but more narrowly, to describe the process of contamination by prior information (Wilson & Brekke, 1994).

Whereas the original heuristics and biases program focused on situations where only heuristics were evoked, and the conjunction fallacy paper examined how heuristics and rules might compete, Griffin & Tversky (1992) described how strength of impression and the statistical weight of evidence might combine. Using the anchoring and adjustment process as the "master heuristic," they suggested that people typically anchor on the strength of their impressions and then adjust (insufficiently) according to rule based arguments about sample sizes or evidential validity. In "support theory," Tversky and his colleagues (e.g. Tversky & Koehler, 1994; Rottenstreich & Tversky, 1997) further developed the notion that even when relatively high-level controlled processes are used to combine and evaluate evidence, intuitive probability judgments are heuristic rather than statistical. In this model, intuitive probability corresponds to an assessment of the relative balance of support for and against a hypothesis. The support may come from direct heuristic assessments or the combination of heuristic assessments and logical arguments. But the use of support as probability is fundamentally non-statistical in two ways. First, the combination of evidence used to derive support need not follow Bayesian rules. Both judgmental heuristics and abstract rules can be used as arguments and summarized by a simple evidence operator, rather than combined according to a Bayesian-type algorithm. Consistent with this, rated balance of support for a hypothesis (an assessment of what is available in the judge's *mind*) is virtually a perfect predictor of the judged probability of the event (the outcome in the *world*), even though the evidence available to the judge cannot be exhaustive (Koehler, 1996; Brenner, 1995). Second, because the relative support is derived according to a *specific hypothesis*, it will change when the objectively identical event is specified in different ways. For example, the judged probability of death due to homicide increases when it is "unpacked" into homicide by an acquaintance or homicide by a stranger (Rottenstreich & Tversky, 1997). Thus, from the high-level evaluation of evidence to the low-level natural assessments, the message from the heuristics and biases program is that intuitive probability judgment is based on "heuristics all the way down."

Heuristic Models and Models of Heuristics

In this section we briefly sketch a framework that we are in the process of developing to model the view of judgmental heuristics as developed in the "second wave." We distinguish five roughly sequential types of variables that progress from features of the environment, to automatic assessments, to basic heuristics, to analytic arguments, to combination heuristics (see Figure 1). Note that not all judgments will entail a mental journey through all levels; processing will stop at the basic heuristics if there is nothing available to combine with the heuristic output.

As Figure 1 shows, we distinguish between three types of automatic assessments that underlie the basic judgmental heuristics: *object based computations* such as similarity or



Note: Boxes on left (underlined words) refer to features of the environment (sources of influence); boxes to the right (italic text) refer to initial low-level automatic processes prompted by stimulus and question; ovals (bold text) refer to semi-automatic heuristic operations that are activated to answer questions; boxes to the farther right (capital text) refer to analytic conscious processes that require cognitive effort.

Figure 1 A sketch of a model of judgmental heuristics causal relatedness, *process monitoring operations* that assess qualities such as mental ease versus effort, and *biased consideration* of information due to prior questions and comparisons. Object based computations tend to fire automatically in response to the stimulus field and are directly translated onto the judgment scale by "direct heuristics" such as representativeness. Process monitoring operations are triggered by specific questions, and their outputs are used by "attributional heuristics" unless discounting cues are present (e.g. Schwarz & Clore, 1988; Strack, 1992). Notable examples of attributional heuristics include the "fluency heuristic" (e.g. Kelley & Jacoby, 1998), in which people use the ease with which items are read or recognized to infer familiarity; using "feelings as information" (Schwarz & Clore, 1988) to infer general well-being from momentary mood state; and recent work on question based availability effects (Schwarz, 1998). These are all instances of attributional heuristics because the use of process monitoring can be blocked or modified if the information is attributed to the environment (i.e. familiarity due to a prior exposure, feelings due to a situational influence, and ease due to the demand to generate a certain number of examples).

A prior question or the presence of a salient value can trigger the third type of automatic assessment, the intentional or unintentional consideration of a hypothesis. In many circumstances merely considering a hypothesis leads to anchoring biases, such that the answer to a following question is contaminated by the initial question or suggestion (Wilson, Houston, Etling, & Brekke, 1996). Note that modern conceptions of anchoring (e.g. Kahneman, 1992) describe it as a bias rather than a heuristic because the initial consideration of a value or hypothesis is rarely *chosen* to serve as an aid in answering the following question. Anchoring effects can be discounted using external cues, but only under very special circumstances (Mussweiler & Strack, 1999).

The final set of heuristics, those involved in evidence combination and evaluation, receives holistic impression based input from the direct heuristics and attributional heuristics, and combines them with whatever rule based or analytic information is salient. These "higher" level heuristic processes operate on consciously represented propositions, combining impressions and rule based arguments, and operate through "inferential" or "argument based" heuristics such as anchoring and adjustment and averaging. Notable rule based arguments include the "statistical heuristics" identified by Nisbett and colleagues (e.g. Nisbett, Krantz, Jepson, & Kunda, 1983); these are simple but abstract rules of thumb such as "you can't learn much from a single experience – it may have been a fluke." Statistical heuristics are most likely to be used in contexts (such as gambling or sports) where statistical or sampling considerations are most salient. The term "inferential heuristics" indicates that even statistical principles are treated as arguments rather than as computational guidelines, except when statistical theory is formally invoked to force the judge to follow optimal algorithmic rules.

According to this model, heuristic processes can be arranged along a continuum ranging from purely impression-based (and the result of purely automatic processes) to purely argument-based (and the result of purely controlled processes). A rough ordering of heuristic processes runs from direct evaluation of impression strength (where the impression *is* the judgment, as in the Tom W. problem), to prediction by evaluation (where the impression is translated into an output scale of judgment, as in non-regressive prediction based on an interview), to anchoring and adjustment via biased accessibility (where a preliminary comparison primes a biased set of evidence that gives rise to a biased judgment: see chapter 11, this volume), to anchoring and adjustment via biased combination (where an existing, perhaps irrelevant, value is adjusted according to an impression), to attributional discounting (where the impression is "checked" for relevance and unbiasedness: Strack, 1992), to argument evaluation (where specific rules or arguments are consciously considered). The extent to which impressions versus arguments are used depends primarily on the transparency of the problem and secondarily on the resources and motivation of the judge. That is, for highly opaque problems such as the "social judgment" problems favored by Kahneman and Tversky, only impression based (perceptual) heuristic processes will be used, regardless of whether cognitive effort and motivation are high or low.

Although our model of heuristics includes automatic processes (natural assessments), conditionally automatic processes (process monitoring that is triggered by attention to particular questions), and controlled processes (attributional inferences and rules of reasoning), it provides a very different perspective than current dual-processing models of social cognition. Dual-process models build on the cognitive miser metaphor by postulating that people typically use low-effort strategies to process information and make decisions, especially when their cognitive resources are strained, but when highly motivated and/or involved, people are capable of using qualitatively different high-effort strategies. In the heuristic–systematic model of persuasion (e.g. Chaiken, Liberman, & Eagly, 1989), for example, depending on motivation, involvement, and cognitive resources, people respond to persuasive messages either with heuristic processing or with systematic processing. Heuristic processing consists of using "if . . . then" rules that operate only on the "shallow" surface structure of the information presented, such as "If it's a credible source, accept the message"; systematic processing consists of deeper processing of the actual meaning of the message. The "stereotypes as heuristics" model (Bodenhausen, 1993) is similar, with variables such as time pressure and mood determining whether stereotypic beliefs are used to guide judgment or whether the available information is deeply processed. Note that depth of processing is simply not *relevant* to the standard heuristics and biases model; in general, the heuristics and biases model describes judgment in high motivation and high capacity contexts, where heuristic processing is dominant simply because of its direct, perceptual nature. However, we would expect the contribution of statistical rules and attributional discounting to be reduced under conditions of high cognitive load, time pressure, or low motivation (e.g. Schwarz, 1998).

According to this view, direct heuristics are not strategically employed to avoid the use of more deliberate rules of reasoning, but are unavoidable aspects of human thought. Rule-based reasoning requires that the heuristic output is "overruled" by deliberate strategies at the level of heuristic evaluation of evidence, and both the recruitment and power of rules will vary according to the context. In this account, motivation and incentives should be less effective in producing rule based thinking than content domain (e.g. strict chance set-ups versus social judgments) and problem structure. This is supported by a recent study (Stephan, 1998) on advanced business students and stock-market professionals. He found that neither incentives, nor domain specific expertise, nor need for cognition (Petty & Caccioppo, 1986) substantially weakened the effects of anchoring, the gambler's fallacy, or the conjunction fallacy. Similarly, Lerner & Tetlock (1999) concluded that the extra effort and motivation induced by accountability serve to increase bias as commonly as to diminish bias. Note that violations of rationality due to ideological and emotional influences are not inconsistent with the heuristics and biases approach, but are simply treated as additional sources of bias (or in some cases, as additional sources of rule based thinking).

Conclusion

In this chapter we have telescoped thirty years of ground-breaking and controversial research into only a few pages. We have been able to highlight only a small selection of the vast literature in this area, and because we have only been able to include a few research examples in the Appendix, much of the richness has been lost. Nonetheless, there is value in the sweeping historical panorama, despite the details and subtleties of the landscape that are lost to sight. The heuristics and biases approach has enriched social psychological theory and research, and in turn, social psychologists have helped enrich the heuristics and biases perspective.

Appendix: A Few Classic Demonstrations of Heuristics and Biases

1 "Tom W.:" Representativeness, non-regressive predictability, and base rate neglect

Tom W. is of high intelligence, although lacking in true creativity. He has a need for order and clarity, and for neat and tidy systems in which every detail finds its appropriate place. His writing is rather dull and mechanical, occasionally enlivened by somewhat corny puns and by flashes of imagination of the sci-fi type. He has a strong drive for competence. He seems to have little feel and little sympathy for other people and does not enjoy interacting with others. Self-centered, he nonetheless has a deep moral sense.

A *similarity* group ranked nine areas of graduate study in terms of "how similar is Tom W. to the typical graduate student." A *prediction* group was informed that the description was written by a high school psychologist on the basis of projective tests, and then ranked the nine areas of graduate study in terms of "the likelihood that Tom W. is now a graduate student in each of these fields. " A *base-rate* group estimated the percentage of graduate students in each field without reading the description.

Results: across the nine graduate fields (e.g. engineering, social sciences, business administration), the judged likelihood correlated .97 with ranked similarity, but – .65 with estimated base rate (Kahneman & Tversky, 1973).

2 "Jack:" Representativeness and base rate neglect

A panel of psychologists have interviewed and administered personality tests to 30 (70) engineers and 70 (30) lawyers. . . . You will find on your forms five descriptions, chosen at random from the 100 available descriptions. For each description, please indicate your probability that the person described is an engineer, on a scale from 0 to 100.

Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles.

Results: when given no information about the individual, respondents predicted the base rate likelihood (either 30 percent or 70 percent depending on condition). Respondents typically gave a 80–90 percent probability that Jack was an engineer, regardless of whether they were presented with the 30 percent base rate or 70 percent base rate conditions (Kahneman & Tversky, 1973). This holds true even when respondents sample the description from an urn, and estimate the base rates for themselves (Griffin & Buehler, 1999).

3 *"Replicating a study:" Representativeness and sample size neglect*

Suppose you have run an experiment on 20 Ss, and have obtained a significant result which confirms your theory ($z = 2.23$, $p < .05$, two-tailed). You now have cause to run an additional group of 10 Ss. What do you think the probability is that the results will be significant, by a one-tailed test, separately for this group?

Results: expert mathematical psychologists estimated a median probability of .85 that the replication would be "significantly" successful. In fact, the normative value is slightly less than .50 (Tversky & Kahneman, 1971). This is presumably because people expect a sample to be overly "representative" of the population, and so underestimate sampling variability.

4 *"Blue cab–green cab:" Causal vs. incidental base rates*

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85 percent of the cabs in the city are Green and 15 percent are Blue (or)

(a′) Although the two companies are roughly equal in size, 85 percent of cab accidents in the city involve Green cabs and 15 percent involve Blue cabs.

(b) A witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80 percent of the time and failed 20 percent of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?

Results: when the "incidental" base rates are presented, as in (a), the median and modal answers are .80, demonstrating base rate neglect. When the "causal" version is presented, as in (a′), the median answer was .60, demonstrating an effect of the base rates (Tversky & Kahneman, 1982). This difference presumably occurred because the causal base rate, implying a certain level of carelessness or accident proneness, provided the Green (but not the Blue) cab company with a "causal disposition" to have accidents (Kahneman & Varey, 1991).

5 *"Linda:" Representativeness, the conjunction fallacy and the neglect of set inclusion (extensionality)*

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in an antinuclear demonstration.

Please rank the following statements by their probability.

Linda is a bank teller.
Linda is a bank teller and is active in the feminist movement.

Results: when only two alternatives were presented, making the problem "transparent" to those who knew the conjunction rule of probability, only statistically naive students rated the conjunction as more probable than the simple event. However, when the two key phrases are embedded in a set of eight targets, even statistically sophisticated students showed a massive conjunction effect (85 percent rated the conjunction more probable; Tversky & Kahneman, 1982, 1983).

6 *"Invasion:" Causality, the conjunction fallacy, and scenario-based prediction*

Please evaluate the probability of (either):

A complete suspension of diplomatic relations between the USA and the Soviet Union, sometime in 1983. (or)
A Russian invasion of Poland, and a complete suspension of diplomatic relations between the USA and the Soviet Union, sometime in 1983.

Results: professional analysts at a forecasting conference rated the second (conjunctive) version as significantly more probable (0.47 percent) than the first (simple event) version (0.14 percent) (Tversky & Kahneman, 1983).

7 *"Heart attacks:" Frequency, representativeness, and the conjunction fallacy*

A health survey was conducted in a sample of 100 adult males in British Columbia, of all ages and occupations. Please give us your best estimate of the following values:

How many of the 100 participants have had one or more heart attacks?
How many of the 100 participants both are over 55 years old and have had one or more heart attacks?

Results: only 25 percent of the students surveyed in this frequentist version judged that the second (conjunctive) category was more numerous compared to about 65 percent surveyed in a percentage version. Asking participants to estimate the number of men over 55 years old in the sample further reduced the incidence of the conjunction fallacy to 11 percent (Tversky & Kahneman, 1983).

References

Adams, J. K., & Adams P. A. (1961). Realism in confidence judgments. *Psychological Review, 68*, 33–45.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35*, 303–314.

Anderson, N. H. (1996). Cognitive algebra versus representativeness heuristic. *Behavioral and Brain Sciences, 19*, 17.

Asch, S. E. (1952). *Social psychology*. New York: Prentice-Hall.

Bar-Hillel, M. (1979). The role of sample size in sample evaluation. *Organizational Behavior and Human Decision Processes, 24*, 245–257.

Bar-Hillel, M., & Neter, E. (1993). How alike is it versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology, 65*, 1119–1131.

Barone, D. F., Maddux, J. E., & Snyder, C. E. (1997). *Social cognitive psychology*. New York: Plenum Press.

Bellhouse, D. R., & Franklin, J. (1997). The language of chance. *International Statistical Review, 65*, 73–85.

Bodenhausen, G.V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping*. San Diego: Academic Press.

Brenner, L. A. (1995). A stochastic model of the calibration of subjective probabilities. Unpublished doctoral dissertation, Stanford University.

Brenner, L. A., Koehler, D. J., Liberman, V., & Tversky, A. (1996). Overconfidence in probability and frequency judgments: A critical examination. *Organizational Behavior and Human Decision Processes, 65*, 212–219.

Camerer, C. (1987). Do biases in probability judgment matter in markets? Experimental evidence. *The American Economic Review, 77*, 981–998.

Campbell, D. T., & Kenny, D. A. (1999). Primer on regression artifacts. New York: Guilford.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 215–252). New York: Guilford Press.

Christensen-Szalanski, J. J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgment and decision literature. *American Psychologist, 39*, 75–78.

Cohen, J. (1960). *Chance, skill, and luck*. Baltimore: Penguin Books.

Cohen, J. (1972). *Psychological probability*. London: George Allen.

Cohen, J., & Hansel, C. E. M. (1956). *Risk and gambling*. New York: Philosophical Library.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4*, 317–370.

Cohen, M. S. (1993). The naturalistic basis of decision biases. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 51–102). Norwood, NJ: Ablex.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment and uncertainty. *Cognition, 58*, 1–73.

Dawes, R. M., Mirels, H. L., Gold, E., & Doahue, E. (1993). Equating inverse probabilities in implicit personality judgments. *Psychological Science, 4*, 396–400.

De Finetti, B. (1970). Logical foundations and measurement of subjective probability. *Acta Psychologica, 34,* 129–145.

Dougherty, M., Gettys, C., & Ogden, E. (1999). MINERVA–DM: A memory processes model for judgments of likelihood. *Psychological Review, 106,* 180–209.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York: Wiley.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review, 103,* 193–214.

Fischoff, B., & Bar-Hillel, M. (1984). Focusing techniques: a shortcut to improving probability judgments? *Organizational Behavior and Human Decision Processes, 34,* 175–194.

Fischoff, B., & Beyth-Marom, R. (1988). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90,* 239–260.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond heuristics and biases. *European Review of Social Psychology, 2,* 83–115.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York, Wiley.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103,* 592–596.

Gigerenzer, G. (1998). Ecological intelligence: An adaptation for frequencies. In D. Dellarosa Cummins & C. Allen (Eds.), *The evolution of mind.* New York: Oxford University Press.

Gigerenzer, G., & Murray, D. J. (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Lawrence Erlbaum.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 513–525.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98,* 506–528.

Gilovich, T., & Savitsky, K. (1999). Like goes with like: The role of representativeness in everyday beliefs. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *The psychology of judgment: Heuristics and biases.* New York: Cambridge University Press.

Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior and Organizations, 17,* 31–57.

Griffin, D. W., & Buehler, R. (1999). Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology, 38,* 48–78.

Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference and human misunderstanding. In M. P. Zanna (Ed.), *Advances in experimental social psychology* Vol. 21. New York: Academic Press.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology, 24,* 411–435.

Griffin, D. W., & Varey, C. A. (1999). *Frequency, probability, and aggregation.* Unpublished manuscript, University of Sussex.

Heath, L., Bryant, F. B., Edwar, J., Fugres-Bakazar, J., Henderson-King, E., Posavacs, E., & Tinsdale, R. S. (Eds.) (1994) *Social psychological applications to social issues, III: Applications of heuristics and biases to social issues.* Chicago: Loyola University Press

Jones, S. K., Jones, K. T., & Frisch, D. (1995). Biases of probability assessment: A comparison of frequency and single-case judgments. *Organizational Behavior and Human Decision Processes 61,* 109–122.

Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes, 51,* 296–312.

Kahneman, D. (1998). Rationality, utility, and the mind. Lionel Robbins Memorial Lecture, London School of Economics.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153.

Kahneman, D., & Tversky, A. (1972). Subjective probability: a judgement of representativeness. *Cognitive Psychology, 3,* 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in Management Science, 12,* 313–327.

Kahneman, D., & Tversky, A. (1982a). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–210). Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (1982b). Variants of uncertainty. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 509–520). Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (1982c). On the study of statistical intuitions. *Cognition, 11,* 123–141.

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103,* 582–591.

Kahneman, D., & Varey, C. A. (1990). Propensities and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology, 59,* 1101–1110.

Kahneman, D., & Varey, C. A. (1991). Mental contamination. Paper presented at the annual meeting of the Society for Experimental Social Psychology, Buffalo.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.), *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feeling. *Acta Psychologica, 98,* 127–140.

Keynes, J. M. (1921). *A treatise on probability.* London: MacMillan.

Kleiter, G. D., Krebs, M., Doherty, M. E., Garavan, H., Chadwick, R., & Brake, G. (1997). Do subjects understand base rates? *Organizational Behavior and Human Decision Processes, 72,* 25–61.

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

Koehler, D. J., Brenner, L., & Griffin, D. (1999). The calibration of probabilities in theory and practice. Unpublished manuscript, University of Waterloo, Waterloo, Ontario.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19,* 1–53.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125,* 255–275.

Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, UK: Cambridge University Press.

Macchi, L. (1995). Pragmatic aspects of the base-rate fallacy. *Quarterly Journal of Experimental Psychology, 48A,* 188–207.

May, R. M. (1986). Inferences, subjective probability, and frequency of correct answers: A cognitive approach to the overconfidence phenomenon. In B. Brehmer, H. Jungermann, P. Lourens, & A. Sevoaan (Eds.), *New directions in research on decision making* (pp. 175–189). Amsterdam: North-Holland.

Meehl, P. E. (1954). *Clinical versus statistical prediction.* Minneapolis: University of Minnesota Press.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficacy of psychometric signs, patterns or cutting scores. *Psychological Bulletin, 52,* 194–216.

Moser, D. V. (1992). Does memory affect judgment? Self-generated versus recall memory measures. *Journal of Personality and Social Psychology, 62,* 555–563.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35,* 136–164.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983) The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate. An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes, 39,* 84–97.

Novemsky, N., & Kronzon, S. (1999). How are base-rates used, when they are used: A comparison of additive and Bayesian models of base-rate use. *Journal of Behavioral Decision Making, 12,* 55–69.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Clinical and Consulting Psychology, 29,* 261–265.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.

Pollack, I., & Decker, I. (1958). Confidence ratings, message reception, and the receiver operating characteristic. *Journal of the Acoustical Society of America, 30,* 286–292.

Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology, 42,* 593–607.

Quattrone, G. A. (1984). Anchors aweigh. Unpublished manuscript, Department of Psychology, Stanford University.

Reeves, T., & Lockhart, R. S. (1993). Distributional versus singular approaches to probability and errors in probabilistic reasoning. *Journal of Experimental Psychology: General, 122,* 207–226.

Robins, R. W., & Craik, K. H. (1993). Is there a citation bias in the judgment and decision making literature? *Organizational Behavior and Human Decision Processes, 54,* 225–244.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology, 37,* 322–336.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review, 104,* 406–415.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Schwarz, N. (1998). Accessible content and accessibility experience: The interplay of declarative and experiential information in judgment. *Personality and Social Psychological Review, 2,* 87–99.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? Informative functions of affective states. In K. Fielder & J. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 44–62). Toronto: Hogrefe.

Schwarz, N., & Sudman, N. (1992). *Context effects in social and psychological research.* New York: Springer-Verlag.

Schwarz, N., Strack, F., Hilton, D., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of "irrelevant" information. *Social Cognition, 9,* 67–84.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61,* 195–202.

Shedler, J., & Manis, M. (1986). Can the availability heuristic explain vividness effects? *Journal of Personality and Social Psychology, 51,* 26–36.

Sherman, S. J., & Corty, E. (1984). Cognitive heuristics. In R. S. Wyer & T. Srull (Eds.), *Handbook of Social Cognition,* Vo. I. Hillsdale: Erlbaum.

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effects of ease of imagery. *Personality and Social Psychology Bulletin, 11,* 118–127.

Simon, H. (1957). *Models of man: Social and rational.* New York: Wiley.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 463–489). Cambridge, New York: Cambridge University Press.

Smith, E. (1996). What do connectionism and social psychology have to offer each other? *Journal of Personality and Social Psychology, 70,* 893–912.

Stephan, S. (1998). Anchoring and adjustment in economic forecasts. Paper presented at the ESRC conference on Judgmental Inputs to the Forecasting Process, London.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249–276). Hillsdale, NJ: Erlbaum.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73*, 437–446.

Thomas, G., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*, 295–314.

Tversky, A., & Kahneman, D. (1971). The belief in the "law of small numbers." *Psychological Bulletin, 76*, 105–110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review, 101*, 547–567.

Varey, C. A. (1991). Crosstalk and contamination in cognitive processes. Unpublished dissertation, University of California, Berkeley.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116*, 117–142.

Wilson, T. D., Houston, C., Etling, K. M., and Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 4, 387–402.

Yates, J. F., & Estin, P. A. (1999). Understanding and improving "people judgment. " Manuscript submitted for publication. University of Michigan, Ann Arbor.

Yates, J. F., Lee, J., Shinotsuka, H., Patalano, A. L., & Sieck, W. R. (1998). Cross-cultural variations in probability judgment accuracy: Beyond general knowledge overconfidence? *Organizational Behavior & Human Decision Processes, 74*, 89–117.

Zukier, H., & Pepitone, A. (1984). Social roles and strategies in prediction: Some determinants of the use of base-rate information. *Journal of Personality and Social Psychology, 47*, 349–360.