

Judgmental Heuristics: A Historical Overview

Dale W. Griffin, Richard Gonzalez, Derek J. Koehler, and Thomas Gilovich

Abstract

The Heuristics and Biases approach to judgment under uncertainty began 40 years ago with the publication of a study of the statistical foibles on the part of research psychologists and statisticians (Tversky & Kahneman, 1971). Since then, this research program has substantially influenced the working assumptions of psychologists and economists about the role of normative models of probability judgment and decision making, while providing a new language of judgmental heuristics. We provide a historical overview of the Heuristics and Biases research program that describes its intellectual antecedents and the special role of the rational actor model in shaping its methods, and we review the program's evolution over the course of three waves of research and theory development.

Key Words: judgment, heuristics, biases, rationality, subjective probability, representativeness, availability, anchoring, adjustment

"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 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, when we consider such matters—even *briefly*—a feeling of certainty or uncertainty seems to "pop out." For example, when a respected British politician was asked whether Kosovo peace talks would lead to a settlement, he stated—with confidence and after only a brief pause—that "the balance of probabilities are 40–60 against." How did he do that?

According to the "Heuristics and Biases" (H&B) approach to human judgment, people typically use cognitive shortcuts that make assessments of likelihood quick and easy but prone to systematic error.

Such shortcuts occur not only in predictions but in retrospective judgments of probability as well, and they can be recognized through signature "biases," as we describe later. Consider a recent article in a major 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 explanation of this strange rendering of the odds based on judgmental heuristics is that people incorrectly (but easily and effortlessly) judge the target sequence of births to be extremely unlikely because the symmetrical pattern of births does not match and is *not representative* 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.

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, 1973; Tversky & Kahneman, 1971, 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 was 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, philosophy, political science, medical decision making, and legal studies.

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 extensions to the original heuristics and biases research. Our presentation is historical in focus and organization, and readers seeking a more focused treatment of recent reconceptualizations are advised to consult Kahneman and Frederick’s (2005) chapter in the previous edition of this Handbook.

The Rational Model

The classical model of rational choice (see Chater & Oaksford, Chapter 2) 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, which provides a behavioral definition and measure of rationality, the “rational actor” assesses the attractiveness of a given option by evaluating the probability of each possible resultant outcome and combining that subjective probability with the subjective utility or personal value of each outcome. The rational economic actor then chooses the best option on the basis of the optimal probability-weighted utility. Economic theories that guide public policy in areas as diverse and important as taxation, environmental safety, stock market and banking regulation, and Social Security rely on the central assumption that individuals and organizations are rational in this sense. Underpinning this model is a series of axioms, or simple rules of logic, that are defined to be both intuitively and formally compelling in their abstract form. This axiomatic foundation provides a series of sharp tests that clearly assess the degree to which observed judgments fit this specific (and widely applied) rational model. The behavioral findings

of Kahneman and Tversky (and many colleagues) question the fundamental assumptions of this normative model by showing how these axiomatic tests fail under well-specified conditions.

There are many events for which it is easy to calculate the “correct” probability (e.g., the chance of drawing 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 less 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% of the time). Note that coherence can be satisfied with respect to purely internal criteria, whereas calibration is specifically defined with respect 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; by this definition, rationality requires only that people follow the rules of subjective probability and evaluate their own preferences consistently.

The most widely used benchmark of the coherence of probability assessment is Bayes’ rule, which has been described as the “master rule” of categorical inference or categorical prediction (see Fischhoff & Beyth-Marom, 1983, for an early psychologically oriented discussion of Bayesian hypothesis testing; also Griffiths, Tenenbaum, & Kemp, Chapter 3). Bayes’ rule defines how to use probability theory to update the probability of 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 reflects 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 prior odds ratio, which reflects 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 to probability, the chance of encountering a given member in one of many random draws).

The strength of inference that can be drawn from a given body 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 given a positive test result, then a prior odds ratio 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.

The use of Bayes' 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' 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, deaths 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—using the principle of coherence—requires subjective probability judgments to follow Bayes' rule.

Historical Antecedents of the Heuristics and Biases Program

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, at least in terms of achieving their goals.

It is worthwhile digging deeper into the Simonian critique, as Simon borrowed the use of the term "heuristic" from computer science and artificial intelligence to describe simplified yet highly efficient human reasoning principles, setting the stage for all future work on heuristics of judgment and decision making. Simon was trained in the field of public administration, and he was originally interested in modeling how bureaucracies worked (a goal more focused on "description" than on "prescription"). The phenomena that Simon and his colleagues observed could be described as "muddling through"—large organizations seemed to operate on simple rules of thumb in an environment in which no one person or department knows everything but somehow everyone knows just enough to produce an adequate outcome. Later, he turned his attention to the psychology of problem solving, with a particular interest in expert judgment. Simon did not build his theories of bounded rationality on specific psychological principles or processes: He explicitly noted that psychological theories of choice processes were not yet sufficiently developed to inform economics. Instead he used general psychological principles to outline some broad, realistic constraints on rational models as models of actual decision making. These general psychological principles reflected the zeitgeist of cognitive psychology at the time, which focused on the limits of memory and attention.

Su
study
ment
theor
peop
satisf
ing,
a stra
look
an a
far a
fied
vagu
is ar
He
chor
ing
still
S
two
to e
moc
hun
orga
nati
and
erec
pro
late
had
rati
inv
as t
ma
I
196
mo
to
pro
cor
(th
rati
jud
clu
dar
by
toc
im
ser
thu
fou
lov

Simon's realistic constraints set the stage for the study of judgmental heuristics and the field of judgment and decision making more generally. In his theories of bounded rationality, he asserted that people simplify the choice process by searching for a satisfactory rather than an optimal outcome. *Satisficing*, he argued, generally consists of three elements: a strategy that examines local or easy options before looking further afield, a stopping rule that specifies an *aspiration level* that must be met and hence how far afield the search should continue, and a simplified assessment of future value that provides a rather vague clue as to the actual value of the choice. There is another less well-known side to Simon's critique: He also maintained that such simplified methods of choice can do surprisingly well relative to optimizing methods and that "bounded rationality" could still be evolutionarily successful.

Simon offered the field of economics (at least) two other familiar psychological insights that were to echo repeatedly in the development of behavioral models of judgment and decision making. First, the human mind (as well as the aggregate mind of the organization) can only hold on to two or three alternatives at one time. Second, attention is a precious and costly commodity, a fact that must be considered in any description of how judgment and choice processes actually operate. Thus, in the vocabulary later introduced by Kahneman and Tversky, Simon had both a negative agenda (explaining how ideal, rational models were unrealistic and descriptively invalid) and a positive agenda (providing guidelines as to how humans—and animals—might actually make highly sensible, if simplified, choices).

Research by Ward Edwards (reviewed in Edwards, 1968) was designed to test rationality assumptions more directly. Using bags full of colored poker chips to explore how people revised, or "updated," their probabilities in the face of new evidence, Edwards concluded that people are not always well calibrated (that is, their probability judgments are not accurate but biased) but are generally coherent in their judgments. In particular, he and his colleagues concluded that *in general* people do reason in accordance with the rules of probability (as summarized by Bayes' rule); however, 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 main predecessor of the H&B approach. However, there were several other flourishing research programs on subjective probability in the 1950s and 1960s that cast further doubt on the rationality assumption. For example, Adams and 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 to general knowledge questions was too low to justify the judged probability of being correct. Researchers using the Signal Detection model (which also has a Bayesian foundation) to study human perceptual judgments (e.g., Pollack & Decker, 1958) found that the correspondence between the rated probability of a "signal" being present and its actual probability 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., Tails when tossing 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 in proportion to the base-rate frequency of that event (e.g., if a roulette wheel was designed to end up "red" on 70% of spins, a probability-matching bettor would bet "red" on 70% of the trials, instead of betting "red" on *every* trial, which would maximize the expected number of wins).

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 afterward, 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 account of intuitive probability. The H&B program was not a comprehensive attempt to explain the anomalies that littered the field of human judgment, but naturally it was influenced by what came before. It was an attempt to describe some of the most notable elements of human judgment that Kahneman and Tversky observed 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 they had concluded that people were approximately or boundedly rational, within limits determined by their computational capacity. However, there was considerable evidence that the assumption of calibration was generally untenable, and some evidence from Cohen’s work that the axioms that required 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 not rational at all (at least in the classical “rational actor” sense).

Later critics have argued that the H&B program marked a sudden and arbitrary shift away from prior research on conservatism, which largely upheld the assumption of rationality (e.g., Gigerenzer & Murray, 1987; Lopes, 1991). This criticism is hard to support, for, as we explain later, the H&B model is consistent with conservatism as well as with the other anomalies listed earlier. The H&B 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 and test the *processes* thought to underlie the real phenomena of interest and to illuminate specific tests that could sharply reject the behavioral applicability of the underlying axioms. The phenomena to be explained were judgments in the real world that seemed to be at odds with the dictates of probability theory.

Three Heuristics Explain Many Biases: The First Wave of Research

The Heuristics and Biases program began when Amos Tversky, a mathematical psychologist who had worked with Edwards and others on formal measurement models, described the current state of the Behavioral Decision Theory paradigm circa 1968 to Daniel Kahneman, his colleague in the Psychology Department at Hebrew University. Kahneman found the idea of tinkering with formal models such as SEU to make them fit the accumulating empirical evidence to be an unpromising approach to understanding the psychological processes involved in judgment and choice. Instead, he argued, based on his own research on visual attention and processing, the principles of cognition underlying judgment should follow the principles of perception (cf. Brunswik, 1956). Thus, instead of starting with formal models as the basis of descriptive accounts of judgment and decision making, Kahneman and Tversky started with principles of perception and psychophysics and extended them to the kind of processing necessary to evaluate probabilities and assess subjective values.

This approach immediately suggested a guiding paradigm for research on judgment and decision making: the study of visual illusions. The logic of studying perceptual illusions is that failures of a

system
system
examp
enorm
more
little t
size of
sky, b
the ho
the vis
percei
visual
sion a
judgr
provic
percep
lights
ment
not sa
it uses
we cal
a heur
necess
and d.
The
sky’s
practi
who h
stand
succes
paradi
lem d
proces
purpo
bias o
groun
cesses
emple
subsec
is wor
descri
The
the
as c
on
acc
app
by
clo
bec
are

system are often more diagnostic of the rules the system follows than are its successes. Consider, for example, the moon illusion: The full moon looks enormous as it sits on the horizon, but it appears more modestly sized when high in the sky. There is little to learn from the constancy of the perceived size of the moon along the long arc of the overhead sky, but its illusory magnification when it sits on the horizon provides insight about the way that the visual system uses contextual detail to compute perceived distance and hence perceived size. The visual illusion paradigm, like the cognitive illusion approach patterned on it, does not imply that judgments of size are typically wrong—in fact, it provides a map to those situations when intuitive perceptions are likely to be correct—but it highlights the processes by which perception or judgment is *constructed* from imperfect cues. We would not say that the visual system is “irrational” because it uses environmental cues in a heuristic way; rather, we can conclude that using environmental cues in a heuristic way gives rise—in well-specified, but not necessarily common circumstances—to systematic and diagnostic errors or biases.

Thus, the guiding logic of Kahneman and Tversky’s approach to the study of judgment was in practice the opposite of that championed by Simon, who had urged researchers to seek out and understand the environmental factors that maximized the success of simple processes. The cognitive illusion paradigm seeks out those environments or problem descriptions in which the judgment and choice processes people rely on lead to clear errors. The purpose was not to emphasize the predominance of bias over accuracy, but to find the clearest testing grounds for diagnosing the underlying simple processes or *judgmental heuristics* that people habitually employ. This is an important distinction that many subsequent critiques have failed to appreciate, and it is worth quoting Kahneman and Tversky’s original description of the logic of the H&B paradigm:

The subjective assessment of probability resembles the subjective assessment of physical quantities such as distance or size. These judgments are all based on data of limited validity, which are processed according to heuristic rules. For example, the apparent distance of an object is determined in part by its clarity. The more sharply the object is seen, the closer it appears to be. This rule has some validity, because in any given scene the more distant objects are seen less sharply than nearer objects. However, the

reliance on this rule leads to systematic errors in the estimation of distance. Specifically, distances are often over-estimated when visibility is poor because the contours of objects are blurred. On the other hand, distances are often underestimated when visibility is good because the objects are seen sharply. Thus the reliance on clarity as an indication of distance leads to common biases. Such biases are also found in intuitive judgments of probability. (Tversky & Kahneman, 1974, reprinted in Kahneman, Slovic, & Tversky, 1982, p. 3)

The heuristics that Kahneman and Tversky identified were also suggested by the principles of perceptual psychology, especially the organizing principles of Gestalt psychology (e.g., Koffka, 1935). Gestalt psychology emphasized how the perceptual system effortlessly and without awareness creates whole forms even when the information reaching the receptors is incomplete and indeterminate. According to the H&B approach, these underlying heuristics are not a simplified version of an ideal statistical analysis but something completely different. This constituted a key point of differentiation between the H&B account and others before it: “In his evaluation of evidence, man is apparently not a conservative Bayesian: he is not Bayesian at all” (Kahneman & Tversky, 1972, p. 450). Unfortunately, or so it seems to us, this statement was taken by some to imply that the H&B (hu)man was not simply un-Bayesian, but rather stupid.

In a later phase of their collaborative research, Kahneman and Tversky took the perceptual framework they had used to study probability judgment and used it to illuminate decision making under risk, leading to their most complete and formal model, Prospect Theory (Kahneman & Tversky, 1979), for which Kahneman received the Nobel Prize in Economics in 2002. In this model, fundamental perceptual principles such as comparison levels and adaptation (Helson, 1964), diminishing sensitivity, and the privileged status of pain served as the primitives of a model that once again used specific biases and errors as tools of diagnosis (see LeBoeuf & Shafir, Chapter 16).

It is illuminating to compare the evolutionary implications of Simon’s Bounded Rationality and the H&B approach. For Simon, the guiding evolutionary principle was computational realism (i.e., simplified approximation) that nonetheless was well adapted to fit the information environment. For Kahneman and Tversky, the guiding evolutionary

principle was that existing processes in perceptual analysis were coopted as tools for higher level cognitive processing. Although these tools might work well in many environments, they also lead to signature biases that are endemic to human intuition. In many cases, the biases that Kahneman and Tversky were signals of underlying heuristics were already well known. As noted earlier, Meehl and Rosen (1955) had warned clinicians of the danger of neglecting base rates in psychological diagnoses. In other cases, the biases were identified by informal observation, whether of psychologists who seemed to neglect power and underestimate sample sizes, army officers who neglected regression effects in determining the value of rewards versus punishment, or army selection personnel who maintained their belief in the efficacy of interviews despite statistical evidence to the contrary.

Negative and Positive Aspects of the Heuristics Program

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 or three 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, at times, to the complete neglect of other normatively important information. The second stage (or as we describe it later, 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. Finally, the third stage organized a broader set of heuristic processes under the rubric of a dual-process model of reasoning and judgment.

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 the processes underlying 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 reasoning problem could be made "transparent" enough to allow participants to "see through" to its underlying statistical framework, but they argued that between-subject manipulations in nonchance settings were most informative about how people typically reasoned in everyday life.

The Positive Model: The Original Perceptual Metaphor

Along with the negative message that people do not intuitively follow Bayes' rule, Kahneman and Tversky developed a positive 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' 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 "shortcut" 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 upward or downward according to knowledge that comes to mind." Whether such procedures were meant to be conscious or intentional strategies was not explicitly stated in the original papers.

Each of the operations described by Kahneman and Tversky yields an impression of certainty or uncertainty, but the heuristic operations themselves are unaffected by some of the required inputs to 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 *nonextensional* and *nonstatistical*, 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 the case information 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-Lyer 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)

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, and 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 experienced psychologists expect their experimental replications to be just as significant as the original (significant) studies. 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. However, when children are less clever than their parents or replications yield weaker results than their originals, people invariably seek out causal explanations—ignoring the statistical law of regression that operates whenever predictive relationships are not perfect. 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 (Gilovich, 1991).

AVAILABILITY

Given that there are a lot of Canadian comedians, one can probably think of particular examples very readily. There is merit, then, in turning this around and concluding that if one has an easy time thinking of Canadian comedians, there probably are a lot of them. The logic is generally sound and it constitutes the essence of the *availability* heuristic, or the tendency to use the ease with which one can generate examples as a cue to category size or likelihood. But the “probably” in this inference is important. There can be other reasons why examples of a given category are easy or hard to generate and so availability is not always a reliable guide to actual frequency or probability (Kahneman & Tversky, 1973; Macleod & Campbell, 1992; Rothbart, Fulero, Jensen, Howard, & Birrell, 1978; Tversky & Kahneman, 1973).

Kahneman and Tversky (1973) first demonstrated this in a series of classic experiments. In one, participants were asked whether there are more words that begin with the letter “r” or that have “r” as the third letter. Because it’s easier to generate words that start with “r” (red, rabid, ratatouille...) than words that have an “r” in the third position (...Huron, herald, unreasonable), most participants thought there were more of the former than the latter. In reality, there are three times as many words with an “r” in the third position.

Ross and Sicoly (1979) explored the implications of the availability heuristic for everyday social life. They asked couples to specify their own percentage contribution to various tasks and outcomes

that come with living together—keeping the house clean, maintaining the social calendar, starting arguments, and so on. They predicted that each person’s own contributions would be more salient than their partner’s contributions and so both partners would overestimate their own role. When the estimates made by each member of a couple were summed, they tended to exceed the logical maximum of 100%. This was true, notably, for negative actions (e.g., starting fights) as well as positive actions—evidence that it is the availability heuristic and not self-enhancing motivations that is responsible for this effect.

Norbert Schwarz and his colleagues have shown how the availability heuristic can influence people’s self-assessments and, in so doing, also settled an important conceptual issue that lies at the core of the availability heuristic (Schwarz, Bless, et al., 1991; Schwarz & Vaughn, 2002; see also Gabrielcik & Fazio, 1984). Recall that people are assumed to use *the ease* with which they can generate instances of a given category when making judgments about the category. But note that if instances are easy to generate, one will probably come up with a lot of them. So how can we be sure that people are in fact influenced by the ease with which they generate instances (a metacognitive feature) rather than the *number* of instances they generate (a cognitive feature)? Typically, we can’t. What Schwarz and colleagues did was to disentangle these two, usually intertwined features. In one representative experiment, they asked half their participants to think of times they had been assertive and the other half to think of times they had been unassertive. Some of the participants in each group were asked to think of six examples and the others were asked to think of twelve examples. The required number of instances, six and twelve, were carefully chosen so that thinking of six examples would be easy but thinking of twelve would be a challenge.

This manipulation separates ease of generation (process) from the number of examples generated (content). Those asked to think of twelve examples of their assertiveness (or unassertiveness) will think of more examples than those asked to think of six, but they will have a harder time doing so. What Schwarz and colleagues found was that those asked to think of six examples of their past assertiveness later rated themselves as more assertive than those asked to think of twelve examples. The same pattern held for those asked to think of past examples of unassertiveness. Thus, it is the ease with which

peop
exam
men
asked
unas
of th
head
asked
ness
past
In
had
could
the-t
to l
diffic
signi
list t

REP
A
colu
and
nut
walr
ron-
ney
orga
to b
fied
M
ness
asse
cons
to a
the
are
is te
sky,
Judg
from
the
effe
198
a gi
larit
and
(Ka
Kah
T
judg
esiz

people can recall examples, not the number of examples recalled, that dominates people's judgments. The effect was so strong, in fact, that those asked to come up with twelve examples of their own unassertiveness (and who thus had lots of examples of their failure to be assertive on the top of their heads) rated themselves as *more assertive* than those asked to come up with twelve examples of assertiveness (and who thus had lots of examples of their past assertiveness at the top of their heads.)

In a wry application of this paradigm, Fox (2006) had students list either two or ten ways a course could be improved as part of the standard end-of-the-term course evaluation process. Students asked to list ten possible improvements apparently had difficulty doing so, because they rated the course significantly more favorably than students asked to list two ways to improve.

REPRESENTATIVENESS

A university nutritionist informed readers of her column that a tomato "has four chambers and is red" and that eating tomatoes is good for the heart; a walnut "looks like a little brain" and "we now know that walnuts help develop more than three dozen neuron-transmitters (sic) for brain function;" and kidney beans assist with the healthy functioning of their organ namesake (Jones, 2008). This advice appears to be heavily influenced by a second heuristic identified by Kahneman and Tversky: representativeness.

Making judgments on the basis of representativeness reflects the mind's tendency to automatically assess the similarity between two entities under consideration and to use that assessment as input to a judgment about likelihood. Judgments about the likelihood of an object belonging to a category are powerfully influenced by how similar the object is to the category prototype (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1983). Judgments of the likelihood that an outcome stems from a particular cause are powerfully influenced by the similarity between putative cause and observed effect (Gilovich & Savitsky, 2002; Nisbett & Ross, 1980). Judgments about the likelihood of obtaining a given result are powerfully influenced by the similarity between the features of the imagined result and those of the processes thought to be at work (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1971).

The most compelling way to demonstrate that judgments are "powerfully" influenced by a hypothesized process is to show that they are excessively

influenced. Much of the research on representativeness has therefore sought to show that the heuristic leads people to make judgments that violate clear normative standards. Judging whether a sample is likely to have come from a particular generating process by assessing the similarity between the two, for example, has been shown to give rise to a "law of small numbers," or a tendency to believe, contrary to probability theory, that even small samples should be representative of the populations from which they are drawn (which is true of large samples and is captured in the law of large numbers). The belief in a law of small numbers has been established by studies showing that people (including expert statisticians and psychologists) are excessively confident about the replicability of research findings (Tversky & Kahneman, 1971), have difficulty recognizing or generating random sequences (Falk & Konold, 1997; Gilovich, Vallone, & Tversky, 1985; Wagenaar, 1972), and are overly influenced by the relative proportion of successes and failures, and insufficiently influenced by sample size, in assessments of how confident they can be in a particular hypothesis (Griffin & Tversky, 1992).

The work on representativeness that garnered the most attention and sparked the greatest controversy, however, involved experiments demonstrating that the allure of representativeness can prevent people from utilizing base rates or basic set-inclusion principles when making predictions. In one now-classic study (Kahneman & Tversky, 1973), participants were given the following description of an individual enrolled in graduate school:

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.

One group of participants was asked to rank nine disciplines in terms of how closely Tom W. resembled the typical student in that field. A second group ranked them in terms of the likelihood that Tom was actually enrolled in each. A third group simply estimated the percentage of all graduate students in the United States who were enrolled in each discipline.

There were two critical findings. First, the rankings of the likelihood that Tom W. actually studied each of the disciplines were virtually identical to the rankings of how similar he seemed to the typical student in each field. Participants' assessments of likelihood, in other words, were powerfully influenced by representativeness. Second, the rankings of likelihood did not correspond at all with what the participants knew about the popularity of the different disciplines. Information about the base rate, or the a priori likelihood of Tom being a student in each of different fields, was simply ignored.

Experiments like this sparked a long-running controversy about whether and when people are likely to ignore or underutilize base rates (Cosmides & Tooby, 1996; Gavanski & Hui, 1992; Gigerenzer, 1991; Griffin & Buehler, 1999; Koehler, 1996). The controversy was productive, especially of publications, because it yielded such findings as: people are more likely to utilize base-rate information if it is presented after the information about the individual (Krosnick, Li, & Lehman, 1990), if the base rate is physically instantiated in a sampling paradigm (Gigerenzer, Hell, & Blank, 1988, but see Poulton, 1994, p. 153), and if the base rate is causally related to the to-be-predicted event (Ajzen, 1977; Tversky & Kahneman, 1982). But in an important respect the controversy was misguided because the essential idea being put forward was that people's judgments are powerfully influenced by representativeness, not that people never use, or even typically don't use, base rates. Instead, the Tom W. studies and others like it were existence proofs of the power of representativeness to overwhelm all other considerations in at least some circumstances.

ANCHORING

Suppose someone asks you how long it takes Venus to orbit the sun. You reply that you don't know (few people do), but your interrogator then asks for an estimate. How do you respond? You might think to yourself that Venus is closer than Earth to the sun and so it probably takes fewer than the 365 days it takes the earth to make its orbit. You might then move down from that value of 365 days and estimate that a year on Venus consists of, say, 275 days. (The correct answer is 224.7.)

To respond in this way is to use what Tversky and Kahneman called the anchoring and adjustment heuristic (Tversky & Kahneman, 1974). One starts with a salient or convenient value and adjusts to an estimate that seems right. The most notable

feature of such adjustments is that they tend to be insufficient. In most investigations of such "anchoring effects," the investigators take care to ensure that the respondents know that the anchor value is entirely arbitrary and therefore carries no implication whatsoever about what the right value might be. In the initial demonstration, Tversky and Kahneman (1974) spun a "wheel of fortune" device and then asked participants whether the percentage of African countries in the United Nations is higher or lower than the number that came up. After participants indicated whether they thought it was higher or lower, they were asked to estimate the actual percentage of African countries in the United Nations. What they found was that the transparently arbitrary anchor value significantly influenced participants' responses. Those who confronted larger numbers from the wheel of fortune gave significantly higher estimates than those who confronted lower numbers. Anchoring effects using paradigms like this have been observed in people's evaluation of gambles (Carlson, 1990; Chapman & Johnson, 1999), estimates of risk and uncertainty (Plous, 1989; Wright & Anderson, 1989), perceptions of self-efficacy (Cervone & Peake, 1986), anticipations of future performance (Switzer & Sniezek, 1991), answers to general knowledge questions (e.g., Jacowitz & Kahneman, 1995), and willingness to pay for consumer items (Ariely, Loewenstein, & Prelec, 2003).

As the research on anchoring evolved, comparable effects using all sorts of other paradigms have been observed and it appears that such effects are not always the result of insufficient adjustment. Indeed, probably the fairest reading of the anchoring literature is that there is not one anchoring effect produced by insufficient adjustment, but a family of anchoring effects produced by at least three distinct types of psychological processes (Epley, 2004). Epley and Gilovich (2001, 2004, 2005, 2006) have provided evidence that people do indeed adjust insufficiently from at least some anchor values, particularly those that people generate themselves (like the earlier question about Venus). They have found, for example, that people articulate a process of adjusting from self-generated anchors, and that manipulations that should influence adjustment, but not other potential causes of anchoring, have a significant effect on people's judgments. In particular, people who are incidentally nodding their heads while answering, are cognitively busy, or lack incentives for accurate responding tend to be more influenced by *self-generated* anchor values than those

who are busy, or

Mani
erally be
response
anchori
Kahnem
& Gilov
1974; se
an excep
because
ficient a
This qu
by Tho
weiler, 1
Strack &
most an
accessibi
attempt
investiga
5,000 [t
test whe
5,000 [c
hypothe
2007; S
1986), s
tionately
and Stra
sis by sh
anchor
dispropo
participants
average
value w
associate
those asl
car is hig
sequent
inexpens
Oppo
recently
ited by c
asked or
sippi Ri
and ano
than 15
draw a l
pick. Th
longer to
tial anch
anchor
"short,"

who are incidentally shaking their heads, are not busy, or are given incentives for accuracy.

Manipulations such as these, however, have generally been shown to have no effect on participants' responses in the standard (experimenter-generated) anchoring paradigm pioneered by Tversky and Kahneman (Chapman & Johnson, 1999; Epley & Gilovich, 2001, 2005; Tversky & Kahneman, 1974; see Simmons, Leboeuf, & Nelson, 2010 for an exception). At first glance, this is a bit of a puzzle because it raises the question of why, without insufficient adjustment, anchoring effects would occur. This question has been addressed most extensively by Thomas Mussweiler and Fritz Strack (Mussweiler, 2002; Mussweiler & Strack, 1999, 2000; Strack & Mussweiler, 1997). They maintain that most anchoring effects are the result of the enhanced accessibility of anchor-consistent information. The attempt to answer the initial question posed by the investigator—"Is the Nile longer or shorter than 5,000 [800] miles?"—leads the individual to first test whether the given value is correct—is the Nile 5,000 [or 800] miles long? Because people evaluate hypotheses by attempting to confirm them (Evans, 2007; Snyder & Swann, 1978; Skov & Sherman, 1986), such a search generates evidence disproportionately consistent with the anchor. Mussweiler and Strack (2000) provide support for their analysis by showing that information consistent with the anchor value presented to participants is indeed disproportionately accessible. For example, participants who were asked whether the price of an average German car is higher or lower than a high value were subsequently quick to recognize words associated with expensive cars (Mercedes, BMW); those asked whether the price of an average German car is higher or lower than a modest value were subsequently quick to recognize words associated with inexpensive cars (Volkswagen, Golf).

Oppenheimer, LeBoeuf, and Brewer (2008) have recently shown that the semantic activation elicited by different anchors can be quite general. They asked one group of participants whether the Mississippi River was longer or shorter than 4,800 miles, and another group whether it was longer or shorter than 15 miles. They then asked their participants to draw a line equal to the length of a standard toothpick. Those exposed to the high initial anchor drew longer toothpicks than those exposed to the low initial anchor. This suggests that exposure to the initial anchor activated the general concept of "long" or "short," which influenced their representation (and

production) of a standard toothpick. To test this idea, Oppenheimer and colleagues had participants in a follow-up experiment perform a word completion task after being exposed to high or low anchor values. Participants exposed to the high anchors were more likely to form words connoting bigness (BIG for B_G, LONG for _ONG) than those exposed to the low anchors.

Recent research suggests that there is likely a third source of anchoring effects: pure numeric priming. That is, an anchor activates its own numeric value and those close to it, which are then highly accessible and influential when the person tries to fashion a response. In one notable experiment, participants were asked whether the runway at Hong Kong International Airport was longer or shorter than 7.3 kilometers or 7,300 meters and were then asked to estimate the cost of an unrelated project. Those asked the question in terms of meters gave higher estimates on the second, unrelated task than those asked the question in terms of kilometers—presumably because the latter primed smaller absolute numbers (Wong & Kwong, 2000). Although some have argued otherwise, this does not appear to be the result of the differential accessibility of semantic information consistent with the initial anchor because 7.3 kilometers and 7,300 meters represent the same value, just in different units. More recent research casts further doubt on the possibility that the differential accessibility of anchor-consistent semantic information is responsible for such effects. Critcher and Gilovich (2008) asked participants what percentage of the sales of a P-97 (or P-17) cell phone would be in the European market. Participants estimated a higher percentage of European sales for the P-97 than the P-17. Note that the process that would give rise to the heightened accessibility anchor-consistent semantic information (testing whether the anchor value might be the correct value) is not applicable here. It seems far-fetched to maintain that participants asked themselves whether part of the model label (97 or 17) might be the European market share.

The Negative Model: Normative Neglect and Its Discontents

The "negative" conclusion from this program of research—that intuitive judgments typically reflect only case-specific evidence; that people neglect base rates, evidence diagnosticity, sample size, 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 H&B program, and it was later made explicit by Griffin and Tversky's (1991) "strength-weight" theory and then modeled by Brenner's (1995, 2003) random support theory. Koehler, Brenner, and Griffin (2002) 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.

As noted earlier in the discussion of the base-rate fallacy and the Tom W. study in particular, criticisms of the "neglect" message began soon after the early laboratory studies were published. 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, p. 17). 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. Tversky and Kahneman (1982) accepted that within-subject designs revealed the *capacity* for rule-based thinking, whereas between-subject designs revealed the *actual* application of rules in practice.

Many economists, whose theories would suffer most if the H&B challenge to classical rationality was widely accepted, wondered about whether the various 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 setup (balls sampled from bingo cages), with both sample evidence and base rates determined by drawing balls from a cage, a context that should make the sampling mechanism salient and "transparent," there was still considerable evidence of heuristic thinking. Similarly, studies of business students in market games involving repeated plays and real incentives also revealed biased judgments in accord with the H&B account (Camerer, 1987), but the biases

seemed to decline with repeated playing of the game (see Camerer & Smith, Chapter 18). It is important to note, however, that studies of 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, 1996; Yates et al., 1998).

We next turn to two critiques that have attracted considerable research attention and raise questions about the fundamental underpinnings of the first wave of the H&B research. The first claim is that findings of the program are merely artifacts of the conversational rules between subject and experimenter; the second is that H&B researchers have confused different definitions of probability.

Some commentators have claimed that many of the apparently "irrational" judgments observed in various 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 H&B research discussed later (e.g., Kahneman & Miller, 1986). Kahneman and Tversky (1982b) 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. In that same chapter, they further noted the relevance of Grice's maxims of communication to their problems (see Grice, 1975; Hilton & Slugoski, 2000) and later explicitly attempted to develop judgment tasks that avoided 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 interpretation. 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)
by su
uals.
cide
died
test t
to rea
indiv
matic
used
provi
versat
diffic
why i
to gu
other
ment
Th
on th
studie
neglec
to jud
been
ing is
tings
exam
attent
pants
when
the pe
Strack
attent
were t
sampl
it was
be ten
of rep
world
logic d
implic
leads t
struct
underc
ticipar
engine
on cha
manip
that ju
a cont
heurist
constr
first e

P(H|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” (p. 198). To test this hypothesis, Macchi changed the key phrase to read “1% of married individuals and 3% 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 provided. Of course, it is possible to reapply 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 they 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 and 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) reported greater attention to base-rate information when participants were told that the personality sketch was randomly sampled by a computer than when they were told it was written by a psychologist. Thus, one might be tempted (despite the many other demonstrations of representativeness in the laboratory and the real world) to conclude that the proper 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” sense-making interpretation that does not undercut the H&B position. In both studies participants were presented only with a low base rate of engineers; inferences about base-rate use were based on changes in judgment in a paradigm that did not manipulate base rate. Thus, these studies suggest not that judgment by representativeness is an artifact of a contrived experimental situation, but rather that heuristics operate upon information that is actively constructed by the perceiver. As noted, one of the first examinations of the role of conversational

implicature (Grice, 1975) in the H&B paradigm was by Kahneman and Tversky in the final chapter of the Kahneman, Slovic, and Tversky (1982) book (Kahneman & Tversky, 1982b, p. 502), which marks the boundary between the first and second waves of the H&B research tradition.

The second major critique is the claim that the H&B program is built solely (and narrowly) on questions about probability judgments for unique events. Some 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 in 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 20th 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 away lives, or even throwing away happiness.

The frequentist critics of the H&B 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., Cosmides & Tooby, 1996; Gigerenzer, 1994; 1998; Jones, Jones, & Frisch, 1995). However, proponents of the H&B approach have explored this possibility for some time, in what we term the “second wave” of heuristics research. For example, Kahneman and Tversky (1982b) 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 (including 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 H&B tradition is 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 frequentist formulations—the dispute is about the causes of the discrepancy and its implications for understanding the classic demonstrations of judgmental heuristics and heuristic thinking in real-world applications.

According to the H&B approach, the discrepancy between single-event probability and aggregate frequency judgments occurs because aggregate frequency judgments are less amenable to heuristic 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. Furthermore, comparisons of the two tasks involve irrelevant confounds because the two scales of judgment are rarely psychologically parallel (Griffin & Buehler, 1999). According to H&B’s frequentist critics, a “frequency format” is consistent with the evolved software of the mind, and single-event “subjective” probability judgments are inherently unnatural (Gigerenzer, 1998, 1994). Supporting this perspective is evidence that people are extremely efficient, and seemingly unbiased, at encoding and storing the frequencies of letters and words to which they have been exposed. 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 or frequentist statistics regularly use expressions of subjective probability to describe their beliefs about the world. A series of studies by Sloman and his colleagues has provided convincing evidence that frequentist representations improve probability judgments when and if they lead to more concrete representations of set-inclusion relations (e.g., Barbey & Sloman, 2007; Sloman, Over, Slovak, & Stiebel, 2003).

Heuristics Unbound: Beyond Three Heuristics

As with any initial statement of a theory, the first wave of H&B demonstrations left some empirical anomalies to be explained. One prominent issue 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 led Tversky and Kahneman (1982) to include the computation or assessment of causality or causal propensity (see Cheng & Buehner, Chapter 12) as a basic heuristic operation, 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, 1982). 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 clear and seemed better captured by a more flexible distinction between “holistic” and “analytic” thinking. Furthermore, the initial statements of the H&B approach contained some ambiguity with regard to whether judgmental heuristics were deliberate strategies to avoid mental effort or were 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 perceptual metaphor, based on prediction by evaluation, that contrasted an “inside” and an “outside” perspective on a prediction problem. 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 afterward, 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 that has won 9 of its last 10 games) may induce a computation of a “causal disposition” connected to that team (Kahneman & Varey, 1990a). These two

approach specific a tinguishe evaluated ner (“assc that requitions) bet

A key included (Tversky primarily 1983 papoped the and clari statistical Kahnema 1982a), tinguishe cognitive and spor from exp which are a questio include c and coun

Tversk tion rule between l. They argu ity (no co than eith most basi understoo Thus, in when the junction l as more p

For exa the follow

Bill is unima lifeless but wo

They v eight poss is an acco (3) Bill is Ninety-ty higher rat of the wo (2) and sc

approaches blurred the distinction between case-specific and statistical information, and instead distinguished between information that can be directly evaluated by natural assessments in a holistic manner ("associationist" computations) and information that requires logical inference (rule-based computations) before it can be used.

A key paper in this second wave of research included the exploration of the *conjunction fallacy* (Tversky & Kahneman, 1983). Although cited primarily for the memorable "Linda problem," the 1983 paper on the conjunction fallacy 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 & Miller, 1986; Kahneman & Tversky, 1982a), Kahneman, Tversky, and colleagues distinguished low-level "natural" or routine or basic cognitive assessments that are relatively automatic and spontaneously evoked by the environment, from explicit, higher level judgmental heuristics, which are typically evoked by an attempt to answer a question. Clear candidates for natural assessments include computations of similarity, causal potency, and counterfactual surprise.

Tversky and 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 one of 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.

For example, participants in one study were given the following description of an individual:

Bill is 34 years old. He is intelligent but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

They were then asked to rank the likelihood of eight possible life outcomes for Bill, including (1) Bill is an accountant, (2) Bill plays jazz for a hobby, and (3) Bill is an accountant who plays jazz for a hobby. Ninety-two percent of the respondents assigned a higher rank to (3) than to (2), even though any state of the world that satisfies (3) automatically satisfies (2) and so (3) cannot be more likely than (2).

Because the conjunction fallacy violates one of the most basic rules of probability theory, Kahneman and Tversky (1983) anticipated controversy and provided a wide-ranging discussion of alternative interpretations. They included additional controls for the possibility that respondents misunderstood the words "and" or "or;" they made sure that the same effects occurred with frequencies as well as probabilities and that the effect applied when reasoning about heart attacks as well as when reasoning about personality descriptions; and they made sure that the same effects obtained with expert and seasoned political forecasters as with college students. Nonetheless, the anticipated controversy ensued, centering around participants' interpretation of the conjunction (e.g., Mellers, Hertwig, & Kahneman, 2001), the effects of frequency versus probability response formats (Hertwig & Gigerenzer, 1999), and the limits of laboratory research.

Kahneman and Tversky created conjunctions that "seemed" or "felt" more likely than their constituents by using representativeness (combined events or descriptions were more similar to the target than one or both of the constituents, as in the Bill and Linda examples), 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 findings from 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 nontransparent within-subject designs, and only substantially 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 *and* B must be less than the number of people with A. Except in special circumstances, then, intuitive judgments do not conform to the rules of probability, even when those rules are known and endorsed by the intuitive judges.

Note how this notion is fundamentally different from the "cognitive miser" model of social cognition. Heuristic judgments are not explained as the

result of too little thought due to cognitive laziness or inadequate motivation, but as the result of uncontrolled "thinking too much" in quick and natural ways. This model of spendthrift automatic processes was termed "mental contamination" by Kahneman and Varey (1990b), 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.

Whereas the original H&B program focused on situations in which only heuristics were evoked, and the conjunction fallacy paper examined how heuristics and statistical rules might compete, Griffin and Tversky (1991) described how the strength of impression and the weight of statistical 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 students developed a formal treatment of how perceptions of evidential support are translated into judgments of probability.

Support theory (Rottenstreich & Tversky, 1997; Tversky & Koehler, 1994) was founded, in particular, on earlier observations of systematic violations of extensionality such as the conjunction fallacy and findings from the fault-tree paradigm (Fischhoff, Slovic, & Lichtenstein, 1978). In contrast to probability theory, in which probabilities are assigned to events that obey the laws of set inclusion, in support theory probabilities are assigned to descriptions of events, referred to as *hypotheses*. Support theory thereby allows two different descriptions of the same event to receive different probability judgments, in much the same way that prospect theory accommodated the possibility that different choices might be made to an identical decision depending on how that decision is "framed" or described (Kahneman & Tversky, 1979).

Support theory represents judged probability in terms of the balance of perceived evidential support for a focal hypothesis A and an alternative hypothesis B, such that $P(A,B) = s(A) / [s(A) + s(B)]$. For instance, A might represent the hypothesis that Jack is a lawyer and B the hypothesis that Jack is an engineer. A key feature of support theory is the assumption that hypotheses described at a greater level of detail will tend to have greater perceived support than a hypothesis describing the same event in less detail. For instance, "unpacking" the hypothesis

that Jack is a lawyer into the hypothesis that Jack is either a corporate lawyer, a criminal lawyer, a divorce lawyer, or a tax lawyer tends to increase support and hence, in the earlier example, the judged probability that Jack is a lawyer rather than an engineer. Unpacking the engineer hypothesis, by contrast, is expected to decrease the judged probability that Jack is a lawyer rather than an engineer. Such unpacking effects are particularly likely when the unpacked components are plausible but unlikely to come to mind in evaluating the packed version of the hypothesis; by contrast, unpacking implausible components can actually have the opposite effect, making the unpacked hypothesis seem less likely than its packed counterpart (Slooman, Rottenstreich, Wisniewski, Hadjichristidis, & Fox, 2004).

Support theory offered an overarching account of intuitive probability judgment that could accommodate a variety of heuristic and other reasoning processes by which the judge might evaluate the extent to which a hypothesis is supported by the available evidence. For instance, given a personality sketch of Jack, support for the lawyer hypothesis might be evaluated based on representativeness (i.e., his similarity to a prototypical lawyer). Or, in the absence of personality information, support might be based on the availability of male lawyers in memory. Support for the lawyer hypothesis might even be based on overall frequency or base-rate information. It is notable that the means by which such statistical information is incorporated in the assessment of support need not necessarily follow the specific combination formula of Bayes rule, but it simply serves as an additional argument that may be considered as part of the support assessment process.

In short, support theory continued the development of incorporating a broader set of assessment processes that went beyond basic heuristics and also characterized the role that heuristics play in intuitive probability judgment as a means by which people evaluate how much a body of evidence supports a particular hypothesis. Indeed, many studies using support theory have shown that eliciting ratings of heuristic attributes such as perceived similarity or causal strength and using them as a proxy measure of support in the earlier support theory equation can reproduce intuitive probability judgments quite closely (e.g., Fox, 1999; Koehler, 1996; Tversky & Koehler, 1994). More generally, in terms of the distinction made by Griffin and Tversky, such research has revealed that perceived support typically is highly sensitive to the strength of the available

evidenc
credenc
This sug
heuristic
work its
consider

Along
man an
judgmen
from ot
notable
the "affe
and the
availabil
metacog
ease of
immedia
as an ir
such as
approach
& MacC

AVAILAI

The m
seem mc
winning
that cand
ining wh
seem tha
Cialdini,
effect wa
ability: li
available
what exa
(or not) c
the likeli

Anoth
around t
event is l
fact, men
"fluent."
difficulty
processin
generatin
and fluer
to proces
nitive exp
inference
Dallas, 19
ent name
Kelley, 19
members

evidence and largely insensitive to its weight or credence (e.g., Brenner, Griffin, & Koehler, 2012). This suggests that support is often evaluated in a heuristic manner, though the support theory framework itself can accommodate other, or additional, considerations as well.

Along with these developments, led by Kahneman and Tversky, the second wave of research on judgmental heuristics saw substantial contributions from other cognitive and social psychologists. Two notable extensions included the development of the “affect heuristic” by Paul Slovic and colleagues, and the splitting off of perceptual fluency from the availability heuristic and treating it as an additional metacognitive “natural assessment” along with ease of generation. The affect heuristic uses one’s immediate good/bad affective reactions to stimuli as an input to various judgments and decisions such as valuation, agreement, and more generally, approach and avoidance (Slovic, Finucane, Peters, & MacGregor, 2002).

AVAILABILITY’S CLOSE COUSIN: FLUENCY

The mere act of imagining an outcome can make it seem more likely to occur. Imagining one candidate winning an election makes it seem more likely that that candidate will triumph (Carroll, 1978) and imagining what it would be like to have a disease makes it seem that one is more at risk of getting it (Sherman, Cialdini, Schwartzman, & Reynolds, 1985). This effect was originally interpreted as the result of availability: Imagining the event made it more cognitively available and hence it was judged more likely. But what exactly are the “relevant instances” that easily (or not) come to mind when one is asked to estimate the likelihood of having an ulcer?

Another interpretation of these findings centers around the concept of fluency: Thinking of a target event is likely to have a different feel if one had, in fact, mentally tried it on earlier. It is likely to feel more “fluent.” Fluency refers to the experience of ease or difficulty associated with perception or information processing and is somewhat distinct from the ease of generating instances. A clear image is easy to process and fluent. A phonemically irregular word is hard to process and disfluent. People use the metacognitive experience of fluency as a cue when making inferences about all sorts of judgments (Jacoby & Dallas, 1981; Oppenheimer, 2008). People judge fluent names to be more famous (Jacoby, Woloshyn, & Kelley, 1989), fluent objects to be better category members (Whittlesea & Leboe, 2000), and adages

that rhyme to be more valid than those that don’t (McGlone & Tofiqbakhsh, 2000).

In addition to these direct effects on judgment, fluency appears to influence how people process relevant information. A feeling of disfluency while processing information appears to undermine people’s confidence in what they are doing, leading to something of a “go slow, be careful” approach to judgment and decision making. Thus, people are more likely to choose a default option when choosing between consumer products that are made disfluent (Novemsky, Dhar, Schwarz, & Simonson, 2007). Fluency also appears to influence the level of abstraction at which information is encoded. Given that blurry (disfluent) objects tend to appear to be farther away than distinct objects (Tversky & Kahneman, 1974), one might expect disfluent entities more generally to appear relatively far away. Indeed, Alter and Oppenheimer (2008) found that cities are judged to be farther away when their names are presented in a difficult-to-read font.

A Third Wave: Dual-Process and Two-System Accounts of Judgment Heuristics

As Neisser (1963) noted in an early review of dual modes of cognition, “The psychology of thinking seems to breed dichotomies.” Consistent with this observation, social and cognitive psychologists have recognized that people appear to approach various cognitive tasks in two very different ways (Chaiken & Trope, 1999; Evans, 2004; Kahneman, 2011; Sloman, 1996; Strack & Deutsch, 2004). One involves mental processes that are fast, associationist, and often automatic and uncontrolled. The other involves processes that are slower, rule based, and more deliberate. Scholars in both disciplines have devoted a lot of energy trying to specify the nature of these two types of processes, or “systems” of thought, and to delineate when each is operative and how they interact when people make a judgment or choose a course of action. The two systems have been given many names and, following Stanovich (1999), we refer to them simply as “System 1” and “System 2” for ease of exposition (see Evans, Chapter 8; Stanovich, Chapter 22). (To our knowledge, the term “dual processes” first appeared in Wason & Evans, 1975.)

Given Daniel Kahneman’s long-standing interest in visual attention—instantiated in his classic 1973 book *Attention and Effort*—it is not surprising that the H&B program came to incorporate

both controlled and automatic processes. (In the preface to his 1973 book, Kahneman wrote: "While the allocation of attention is flexible and highly responsive to the intentions of the moment, there are pre-attentive mechanisms that operate autonomously, outside voluntary control...it is easy to notice several aspects or attributes of an object, but it is difficult or impossible to prevent the perceptual analysis of irrelevant attributes" [p. 7].) The 1983 paper on the conjunction fallacy implicitly provided a dual-system analysis of the competition between automatic intuitive processes and effortful rule-based processes, and this analysis was formalized by Sloman (1996) who also drew upon social psychological models of the conflict between a "gut feeling" and a more considered analysis (Denes-Raj & Epstein, 1994; Epstein, 1991).

Perhaps the most striking evidence of two mental systems that guide judgment and behavior is Epstein's work on the "ratio bias" phenomenon (Denes-Raj & Epstein, 1994; Epstein, 1991). Epstein told participants that they could win a prize by blindly selecting a jellybean of a given color from one of two urns. One urn had 1 winning jellybean and nine of another, losing color. The second urn had 9 winning jellybeans and 91 of the losing color. What Epstein found was that many participants chose to select from the larger urn that offered lower odds of winning because they couldn't resist the thought that the larger urn had more winning beans. They did so despite the fact that the chances of winning with each of the urns was explicitly provided for them. When the choice was between a 10% chance in the small urn and a 9% chance in the large urn, 61% of the participants chose the large urn. When it was a contest between 10% in the small urn and 5% in the large urn—odds only half as good in the latter—23% of the participants still chose the large urn.

Epstein attributes this decidedly irrational result to an "experiential" system of reasoning that operates on concrete representations, and hence finds the greater number of winning jellybeans in the large urn to be more promising. This experiential or intuitive impulse, however, usually conflicts with the rational realization that the actual odds are better in the small urn. Some participants explicitly stated that they knew they should pick the smaller urn, but they nonetheless were going with a gut feeling that they were more likely to pick a winner from the large one. This experience of being pulled in two different directions suggests that there are two things—two mental systems—doing the pulling.

This was emphasized by Sloman (1996), who described a possible cognitive architecture consisting of two relatively independent systems to explain the diverse findings implicating dual processes in reasoning, choice, and judgment.

Kahneman and Frederick (2002, 2005; Kahneman, 2011) highlighted these relations between System 1 and System 2 in their influential "third wave" restatement of the H&B program of research. In their "attribute substitution" account, System 1 automatically computes an assessment with some connection to the task at hand—an emotional reaction, a sense of fluency, the similarity between examples or between an example and a category. Both the perceived relevance of the assessment and its immediacy often give rise to the sense that the task is done and that the assessment produced by System 1 *is* the answer being sought. For example, one cause of death is judged to be more common than another because it is easier to think of examples of the former (Slovic, Fischhoff, & Lichtenstein, 1982). One attribute (ease of retrieval) *substitutes* for another, desired attribute (likelihood).

In many circumstances, however, and for a variety of different reasons, System 2 intervenes and deems the automatic assessment inadequate for the task at hand. A more deliberate, rule-based response is given. For example, one might consciously realize, especially if one has received training in statistics and recognizes threats to validity, that a given cause of death is highly available because it is frequently discussed in the media. "In the context of a dual-system view, errors of intuitive judgment raise two questions: 'What features of system 1 created the error?' and 'Why was the error not detected and corrected by system 2?'" (Kahneman & Frederick, 2005, p. 268). The attribute substitution model has captured a great deal of attention because it offered a unified account of a diverse set of judgmental phenomena, such as the role of heuristics and logical rules on probability judgment, happiness assessments, duration neglect in remembered pain, and on contingent valuation methods used to assess people's willingness to pay (WTP) for such things as environmental remediation. As Kahneman and Frederick (2005, p. 287) summarized:

The original goal of the heuristics and biases program was to understand intuitive judgment under uncertainty. Heuristics were described as a collection of disparate cognitive procedures, related only by their common function in a particular judgmental

me
Kei
has
tion
acte
syst
for
dua
pro
that
refl
resp
200
mer
but
can
base

Co
V
trad
tion
reco
imp
tion
the
shoc
exar
base
from
Tu,
mot
isola

domain... It now appears, however, that judgment heuristics are applied in a wide variety of domains and share a common process of *attribute substitution* in which difficult judgments are made by substituting conceptually or semantically related assessments that are simpler and more readily accessible.

The current treatment explicitly addresses the conditions under which intuitive judgments are modified or overridden. Although attribute substitution provides an initial input into many judgments, it need not be the sole basis for them. Initial impressions are often supplemented, moderated or overridden by other considerations, including the recognition of relevant logical rules and the deliberate execution of learned algorithms. The role of these supplemental or alternative inputs depends on characteristics of the judge and the judgment task.

Although the dual-system account of judgmental heuristics is not without its skeptics (e.g., Keren & Schul, 2009), in its very general form it has received broad acceptance. A number of questions, remain, however, about the best way to characterize the operations and interactions of the two systems. Evans (e.g., 2008; see Evans, Chapter 8), for instance, distinguishes default-interventionist dual-process models from parallel-competitive dual-process models. It has been suggested, furthermore, that System 2 should be split into two components reflecting cognitive ability and thinking dispositions, respectively, yielding a triprocess model (Stanovich, 2009). In short, the dual-system approach to judgment under uncertainty has been very influential, but many details will need to be filled in before it can be developed into a comprehensive process-based account of how heuristics operate.

Conclusions

We provided a historical overview of the H&B tradition, its intellectual forebearers, and its evolution through three waves of conceptualization and reconceptualization, but this should not be taken to imply that the program is frozen in the past. In addition to the hearty oak tree of classic H&B research, the program still continues to send out new green shoots of intellectual offspring. One "second-wave" example is a new model of counterfactual reasoning based on a model of semantic evidence that follows from the H&B approach (Miyamoto, Gonzalez, & Tu, 1995). The dual-processing perspective has motivated neuropsychological studies attempting to isolate and locate the brain networks associated with

the "dueling" heuristic and rule-based processes underlying classic H&B demonstrations (e.g., De Neys & Goel, 2011); the impact of experienced versus presented statistical information continues to be an active area of research (e.g., Brenner, Griffin, & Koehler, 2005; Brenner, Griffin, & Koehler, 2012; Fox & Hadar, 2006; Hertwig, Barron, Weber, & Erev, 2004); and the applied impact of the H&B tradition on understanding expert judgment in such fields as finance, political science, law, medicine, and organizational behavior continues to grow (e.g., Koehler, Brenner, & Griffin, 2002; Tetlock, 2005).

Future Directions

Questions to guide future development:

What is the relation between general cognitive ability and susceptibility to heuristic-based judgmental bias?

Are there fundamental individual differences in the tendency to make heuristic-based judgments?

What circumstances facilitate detection and correction of conflict between heuristic and rule-based evaluations?

Are the processes underlying System 2 operations that support "override" of initial, heuristic responses related to more basic inhibitory operations that guide attention in, for example, Stroop, flanker, and go/no-go tasks?

Under what conditions does extended experience in carrying out a particular judgment task reduce susceptibility to base-rate neglect, conjunction errors, and other systematic biases?

Do organizational practices or market interaction consistently attenuate biases associated with use of judgmental heuristics?

Can anchoring and/or adjustment usefully be viewed as the "master heuristic?"

How useful is the distinction between positive versus negative contributions in theory development to psychology more generally?

Acknowledgments

This chapter draws extensively upon reviews presented in Gilovich and Griffin (2010) and Griffin, Gonzalez, and Varey (2001). We acknowledge financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC, Griffin), the Natural Sciences and Engineering Research Council of Canada (NSERC, Koehler), and the National Science Foundation (NSF, Gonzalez, Gilovich).

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.
- Alter, A.L., & Oppenheimer, D. M. (2008). Effects of fluency on psychological distance and mental construal (or why New York is a large city, but *New York* is a civilized jungle). *Psychological Science*, 19, 161–167.
- Anderson, N. H. (1996). Cognitive algebra versus representativeness heuristic. *Behavioral and Brain Sciences*, 19, 17.
- Ariely, D., Loewenstein, G., & Prelec, D. (2003). “Coherent arbitrariness”: Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118, 73–105.
- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Brain and Behavioral Sciences*, 30, 241–298.
- Bar-Hillel, M., (1982). Studies of representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 69–83). New York: Cambridge University Press.
- 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.
- Brenner, L. A. (1995). A stochastic model of the calibration of subjective probabilities. Unpublished Ph.D. dissertation, Stanford University, Palo Alto, CA.
- Brenner, L. A. (2003). A random support model of the calibration of subjective probabilities. *Organizational Behavior and Human Decision Processes*, 90, 87–110.
- Brenner, L., Griffin, D. W., & Koehler, D. J. (2005). Modeling patterns of probability calibration with Random Support Theory: Diagnosing case-based judgment. *Organizational Behavior and Human Decision Processes*, 97, 64–81.
- Brenner, L., Griffin, D. W., & Koehler, D. J. (2012). A case-based model of probability and pricing judgments: Biases in buying and selling uncertainty. *Management Science*. doi: 10.1287/mnsc.1110.1429.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.
- Camerer, C. F. (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). *A primer on regression artifacts*. New York: Guilford.
- Carlson, B. W. (1990). Anchoring and adjustment in judgments under risk. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 665–676.
- Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, 14, 88–96.
- Cervone, D., & Peake, P. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology*, 50, 492–501.
- Chaiken, S., & Trope, Y. (1999) *Dual-process theories in social psychology*. New York: Guilford Press.
- Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation and the construction of value. *Organizational Behavior and Human Decision Processes*, 79, 115–153.
- Cohen, J., & Hansel, C. E. M. (1956). *Risk and gambling*. New York: Philosophical Library.
- Cohen, J. (1960). *Chance, skill, and luck: The psychology of guessing and gambling*. Baltimore, MD: Penguin Books.
- 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.
- Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making*, 21, 241–251.
- Dawes, R. M., Mirels, H. L., Gold, E., & Donahue, 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.
- Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, 66, 819–829.
- De Neys, W., & Goel, V. (2011). Heuristics and biases in the brain: Dual neural pathways for decision making. In O. Vartanian & D. R. Mandel (Eds.), *Neuroscience of decision making* (pp. 125–142). Hove, England: Psychology Press.
- Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York: Wiley.
- Epley, N. (2004). A tale of tuned decks? Anchoring as accessibility and anchoring as adjustment. In D. J. Koehler & N. Harvey (Eds.), *The Blackwell handbook of judgment and decision making* (pp. 240–256). Oxford, England: Blackwell Publishers.
- Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Divergent processing of self-generated and experimenter-provided anchors. *Psychological Science*, 12, 391–396.
- Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin*, 30, 447–460.
- Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally-provided anchors. *Journal of Behavioral Decision Making*, 18, 199–212.
- Epley, N., & Gilovich, T. (2006). The anchoring and adjustment heuristic: Why adjustments are insufficient. *Psychological Science*, 17, 311–318.
- Epstein, S. (1991). Cognitive-experiential self-theory: An integrative theory of personality. In R. Curtis (Ed.), *The self with others: Convergences in psychoanalytic, social, and personality psychology* (pp. 111–137). New York: Guilford Press.
- Evans, J. St. B.T. (2004). History of the dual process theory in reasoning. In K. I. Manktelow & M. C. Chung (Eds.), *Psychology of reasoning: Theoretical and historical perspectives* (pp. 241–266). Hove, England: Psychology Press.
- Evans, J. St. B.T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment*. New York: Psychology Press.
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–278.
- Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104, 301–318.
- Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, 90, 239–260.

Fischhoff,
Sensit
repres
Percep
Fox, C. I
and c
167–1
Fox, C. F
How
Judgm
Fox, C. R
sampl
Barror
Makin
Gabrielcik
estima
and So
Cavanski,
uncert
63, 58
Gigerenzer
pear:
Social
Gigerenzer
probab
(and v
probab
Gigerenzer
tion fo
(Eds.),
Univer
Gigerenzer
and co
Journal
Perform
Gigerenzer
statistic
Gilovich, T
human
Gilovich, T
making
of social
Gilovich, T
role of
beliefs.
Heurist
617–62
Gilovich, T
in bask
Cogniti
Grether, D.
heuristic
Behavior
Grice, H. P.
Morgan
(pp. 225
Griffin, D. V
predictic
Psycholog
Griffin, D. V
tics and l
Tesser &
psycholog

- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representations. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 330-344.
- Fox, C. R. (1999). Strength of evidence, judged probability, and choice under uncertainty. *Cognitive Psychology*, 38, 167-189.
- Fox, C. R. (2006). The availability heuristic in the classroom: How soliciting more criticism can boost your course ratings. *Judgment and Decision Making*, 1, 86-90.
- Fox, C. R., & Hadar, L. (2006). Decisions from experience = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making*, 1, 159-161.
- Gabrielcik, A., & Fazio, R. H. (1984). Priming and frequency estimation: A strict test of the availability heuristic. *Personality and Social Psychology Bulletin*, 10, 85-89.
- Gavanski, I., & Hui, C. (1992). Natural sample spaces and uncertain belief. *Journal of Personality and Social Psychology*, 63, 585-595.
- 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. (1998). Ecological intelligence: An adaptation for frequencies. In D. Dellarosa Cummins & C. Allen (Eds.), *The evolution of mind* (pp. 9-29). New York: Oxford University Press.
- 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., & Murray, D. J. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: The Free Press.
- Gilovich, T., & Griffin, D. W. (2010). Judgment and decision making. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (pp. 542-588). New York: McGraw-Hill.
- Gilovich, T., & Savitsky, K. (2002). Like goes with like: The role of representativeness in erroneous and pseudo-scientific beliefs. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 617-624). New York: Cambridge University Press.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17, 295-314.
- Grether, D. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior and Organization*, 17, 31-57.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics. Volume 3: Speech acts* (pp. 225-242). New York: Seminar Press.
- Griffin, D. W., & Buehler, R. (1999). Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology*, 38, 48-78.
- Griffin, D. W., Gonzalez, R., & Varey, C. A. (2001). The heuristics and biases approach to judgment under uncertainty. In A. Tesser & N. Schwarz (Eds.), *The Blackwell handbook of social psychology* (pp. 207-235). Oxford, England: Blackwell.
- Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 319-356). New York: Academic Press.
- Griffin, D. W., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24, 411-435.
- Helson, H. (1964). *Adaptation-level theory*. New York: Harper.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15, 534-539.
- Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12, 275-306.
- Hilton, D. J., & Slugoski, B. R. (2000). Judgment and decision making in social context: Discourse processes and rational inference. In T. Connolly, H. R. Arkes, & K. R. Hammond (Eds.), *Judgment and decision making: An interdisciplinary reader* (pp. 651-676). Cambridge, England: Cambridge University Press.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology*, 3, 306-340.
- Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, 118, 115-125.
- Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21, 1161-1167.
- Jones, J. (2008, February 12). Truly functional foods. *PressRepublican.com*. Retrieved August 2011, from http://www.pressrepublican.com/0808_health/local_story_042224534.html
- 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. (2011). *Thinking, fast and slow*. New York: Farrar, Strauss, Giroux.
- Kahneman, D., Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge, England: Cambridge University Press.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-293). New York: Cambridge University Press.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136-153.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment 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). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.
- Kahneman, D., & Tversky, A. (1982a). The psychology of preferences. *Scientific American*, 246, 160-173.

- 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, England: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315.
- Kahneman, D., & Varey, C. A. (1990a). Propensities and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology*, 59, 1101–1110.
- Kahneman, D., & Varey, C. A. (October, 1990b). *Mental contamination*. Paper presented at the Annual meeting of the Society for Experimental Social Psychology, Buffalo, NY.
- Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science*, 4, 533–550.
- Keynes, J. M. (1921). *A treatise on probability*. London: MacMillan.
- Koehler, J. J. (1996). The base-rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, 19, 1–53.
- Koehler, D. J., Brenner, L., & Griffin, D. W. (2002). The calibration of expert judgment: Heuristics and biases beyond the laboratory. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 686–715). New York: Cambridge University Press.
- Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World.
- Krosnick, J. A., Li, F., & Lehman, D. R. (1990). Conversational conventions, order of information acquisition, and the effect of base rates and individuating information on social judgments. *Journal of Personality and Social Psychology*, 59, 1140–1152.
- Lopes, L. L. (1991). The rhetoric of irrationality. *Theory and Psychology*, 1, 65–82.
- Macchi, L. (1995). Pragmatic aspects of the base-rate fallacy. *Quarterly Journal of Experimental Psychology*, 48A, 188–207.
- Macleod, C., & Campbell, L. (1992). Accessibility and probability judgments: An experimental evaluation of the availability heuristic. *Journal of Personality and Social Psychology*, 63, 890–902.
- McGlone, M. S., & Tofiqbakhsh, J. (2000). Birds of a feather flock conjointly(?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424–428.
- Meehl, P. E. (1954). *Clinical versus statistical prediction*. Minneapolis: University of Minnesota Press.
- Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 52, 194–216.
- Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, 12, 269–275.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Miyamoto, J., Gonzalez, R., & Tu, S. (1995). Compositional anomalies in the semantics of evidence. *The Psychology of Learning and Motivation*, 32, 319–383.
- Mussweiler, T. (2002). The malleability of anchoring effects. *Experimental Psychology*, 49, 67–72.
- 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.
- Mussweiler, T., & Strack, F. (2000). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology*, 78, 1038–1052.
- Neisser, V. (1963). The multiplicity of thought. *British Journal of Psychology*, 54, 1–14.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Novemsky, N., Dhar, R., Schwarz, N., & Simonson, I. (2007). Preference fluency in choice. *Journal of Marketing Research*, 44, 347–356.
- Oppenheimer, D. M., (2008). The secret life of fluency. *Trends in Cognitive Sciences*, 12, 237–241.
- Oppenheimer, D. M., LeBoeuf, R. A., & Brewer, N. T. (2008). Anchors aweigh: A demonstration of cross-modality anchoring. *Cognition*, 106, 13–26.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Clinical and Consulting Psychology*, 29, 261–265.
- Plous, S. (1989). Thinking the unthinkable: The effect of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology*, 19, 67–91.
- Pollack, I., & Decker, L. R. (1958). Confidence ratings, message reception, and the receiver operating characteristic. *Journal of the Acoustical Society of America*, 30, 286–292.
- Poulton, E. C. (1994). *Behavioral decision theory: A new approach*. Cambridge, England: Cambridge University Press.
- Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, 32, 880–892.
- Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, P. (1978). From individual to group impressions: Availability heuristics in stereotype formation. *Journal of Experimental Social Psychology*, 14, 237–255.
- 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., 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.
- 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., & Vaughn, L. A. (2002). The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 103–119). New York: Cambridge University Press.
- 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 effect of ease of imagery. *Personality and Social Psychology Bulletin*, 11, 118–127.
- Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley.
- Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors? *Journal of Personality and Social Psychology*, 99, 917–932.

Skov, R. B.,
cesses: E
perceive
Social Ps
Slovan, S. i
soning.
Slovan, S. i
Frequen
Behavio
Slovan, S.,
C., & F
superad
Psycholo
Slovic, P., Fi
The affi
Kahnen
intuitiv
Cambri
Slovic, P., F
fears: U
Slovic,
Heurist
Univers
Snyder, M.,
interact
1202–J
Stanovich,
differen
Stanovich,
rithmic
theory?
processe
Univers
Strack, F., &
minant
Review,
Strack, F.,
anchor
Journal
Switzer, F.
motiva
and be
Process
Tetlock, P.
it? Hoo
Press.

- Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22*, 93-121.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3-22.
- Sloman, S. A., Over, D. E., Slovak, L., & Striebel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes, 91*, 296-309.
- Sloman, S., Rottenstreich, Y., Wisniewski, E., Hadjichristidis, C., & Fox, C. (2004). Typical versus atypical unpacking and superadditive probability judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 573-582.
- Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 397-420). Cambridge, England: Cambridge University Press.
- 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). New York: Cambridge University Press.
- Snyder, M., & Swann, W. B. (1978). Hypothesis-testing in social interaction. *Journal of Personality and Social Psychology, 36*, 1202-1212.
- Stanovich, K. E. (1999). *Who is rational? studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.
- Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). Oxford, England: Oxford University Press.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220-247.
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73*, 437-446.
- Switzer, F., & Sniezek, J. A. (1991). Judgment processes in motivation: Anchoring and adjustment effects on judgment and behavior. *Organizational Behavior and Human Decision Processes, 49*, 208-229.
- Tetlock, P. E. (2005). *Expert political judgment: How good is it? How can we know?* Princeton, NJ: Princeton University Press.
- Tversky, A., & Kahneman, K. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*, 105-110.
- Tversky, A., & Kahneman, K. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. (pp. 153-160). 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.
- Wagenaar, W. A. (1972). Generation of random sequences by human subject: A critical survey of literature. *Psychological Bulletin, 77*, 65-72.
- Wason, P. C., & Evans, J. St. B. T. (1975). Dual processes in reasoning? *Cognition, 3*, 141-154.
- Whittlesea, B. W., & Leboe, J. P. (2000). The heuristic basis of remembering and classification: Fluency, generation, and resemblance. *Journal of Experimental Psychology: General, 129*, 84-106.
- Wong, K. F. E., & Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes, 82*, 314-333.
- Wright, W. F., & Anderson, U. (1989). Effects of situation familiarity and financial incentives on use of the anchoring and adjustment heuristic probability assessment. *Organizational Behavior and Human Decision Processes, 44*, 68-82.
- Yates, J. F., & Estin, P. A. (November, 1996). *Training good judgment*. Paper presented at the Annual Meeting of the Society for Judgment and Decision Making, Chicago.
- 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 and 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.