Prejudices about Bias

Baruch Fischhoff

Decision Research
A Branch of Perceptronics
1201 Oak Street
Eugene, Oregon 97401

May 1985

Running Head: Prejudices

Abstract

Much recent research in the area of judgment and decision making has been dominated by documentation of ways in which people's intuitive thought processes can lead them astray. Like other psychological results that have cast doubt on people's abilities, these accounts have generated considerable controversy. From the set of criticisms that have been raised, a set of generic criticisms that could be raised is developed here. These include aspersions of methodological malpractice, advancement of alternative standards of optimality, and development of error theories showing the insensitivity of events to these kinds of problems. Considering these criticisms in general form offers some perspectives on the continuing debate, some guidance on how to improve its productivity, and some hypotheses for analyzing analogous controversies elsewhere in psychology.

Prejudices About Bias

A cynical view of psychologists holds that they make their living observing quirks in other people's behavior. One good quirk, documented through half a dozen articles, is enough to secure tenure in most depart-ments. One megaquirk, à la Freud, may secure a lifelong (or longer) reputation.

Although they may be a road to success, quirks are also a road to controversy. Claims about newly discovered human foibles seem more likely to be argued about than, say, fresh estimates of a particular choice reaction time, threshold of perception, or digit span. Not only do more people feel that they need to express an opinion about the work, but the tenor of those opinions is more likely to become strident than with non-quirk claims making equally fundamental statements about human behavior.

It would take a psychologist to provide a full account of why people worry so about quirks. However, a number of recurrent themes can be identified. Considering them in advance allows those about to claim the existence of a new quirk to anticipate the standard counterclaims, whether to buttress their own findings or to prepare the standard rebut-tals. Considering these themes in the heat of a debate can help sharpen the discussion, saving participants the exercise in lay philosophy needed to cast it in generic terms. Considering the themes in retro-spect may indicate what chance a debate had of being resolved (given the nature of the claims and counterclaims), whether some key concepts or studies might have speeded its resolution, and where the unresolved

conflict has gone.

These themes are discussed here in terms of their emergence in the controversy surrounding the research into judgmental biases associated with Kahneman, Tversky, and their colleagues (or apologists, or fellow travelers, depending on one's prejudices about these biases).[1] Analogous developments might be found in the controversies over Freud's claims of psychodynamic distortions, Piaget's descriptions of children's cognitive limitations, Kuhn's characterizations of the built-in incapacities that come with socialization in a science, as well as other less sensational episodes.

For readers unfamiliar with this research, Table 1 summarizes some of the biases or quirks that investigators in this tradition claim to have identified, along with some pertinent references. The biases are seen as the byproducts of widely used rules of thumb, or heuristics, that people use to direct their judgments under conditions of uncertainty. These rules have some natural appeal and can, at times, be helpful. However, they can also lead to predictable problems when used alone or without qualification. For example, recalling instances of an event occurring can provide a clue to its frequency (the availability heuristic), but may be misleading when its occurrence is unusually memorable or forgettable.

Insert Table 1 about here

"It's Not True"

Psychologists are adept at picking holes in the design and analysis of research. The long apprenticeship that confers this ability is

intended to help them secure the foundations of their own studies and ferret out the justifiable conclusions from studies conducted by others. Although some picking is always in order, it can also be out of order, as when routine criticisms are used to avoid listening to results that one simply does not want to hear. Every colloquium audience has regulars, ready with the same methodological critiques for all comers. Where these standard questions remind investigators of alternative explanations, they serve a useful role. Where they compell discussion of controls that went without saying or potential effects of at best minimal size, then they are tedious. Where they probe at issues that can hardly be answered within the context of the discipline (e.g., how do you know that your subjects <u>really</u> understood the task? How do you know that your subjects weren't trying to make your hypothesis succeed--or fail?), then they can stifle constructive discussion.

A limiting factor on most critiques, even obstructionist ones, is not threatening the shared assumptions of the discipline, those articles of faith without which any interpretation of data would be impossible. For complete radical skepticism, an outsider is often needed. Thus, philosophers can offer (what psychologists might consider) outrageous alternative explanations of what subjects might be thinking in a particular experiment without feeling any compulsion to test the viability of that account for other experiments in the same article, much less seek a single explanation for the full body of evidence. Or, sociologists can refuse to consider results that have not been obtained with proper probability samples of participants. Or, economists can insist that the characterization of the task is incomplete until it includes whatever

factors subjects succeeded in optimizing by their behavior (which is held, by definition, to serve their best interests).

The first of the six sections in Table 2 contains a set of technical questions commonly raised by those who accept the possibility of psychological research into human judgment. Each assumes that something is wrong with the particular procedures that have been used. The claim is that once these instances of methodological malpractice are remedied, the biases will vanish. For example, if the task is made really important to subjects, then they will apply themselves hard enough to get it right.

Insert Table 2 about here

It is a question of philosophy of science whether the burden of proof should lie more with the original investigator to refute these claims or with their proponents to establish their validity. The burden probably shifts toward the latter when they argue not just that the original interpretation is not true, but that an alternative one is. Thus, it is one thing to claim that failure to provide high enough stakes vitiates a claim of apparently biased behavior, and another thing to substantiate the belief that people perform well when they really care about a task.

Given the costs of conducting studies to address every possible counterclaim, there is a temptation to meet skepticism with skepticism. One common expression of the refusal to respond to dataless critiques is arguing that subjects really care about well-staged experiments that treat them respectfully. Another is arguing that experiments give

subjects a better deal than they are likely to get in life, where the analogous problems are typically buried in the confusion of everyday existence, without the opportunity to focus attention on them, to have them neatly circumscribed by the experimenter, and to be rewarded for good judgment (rather than, say, for exuding confidence). A complement of this argument is that it is hard to create extrinsic incentives that will have just the right effect: well-intended reward schemes may inadvertently induce irrelevant competition among subjects, gaming between subject and experimenter, or behavior peculiar to some unexpected interpretation of the particular payoff rule.[2]

Despite the fluency with which such conjecture-refutation interchanges can be conducted and the heat that they are able to generate, some data do get collected. One way to make the data go a little further, and obviate the need for conducting every methodological check on every extant result, is to see whether any general statements can be made about the impact of various methodological controls. A review of all published studies that could be construed as attempts to remedy flaws in tasks that allegedly revealed hindsight bias or overconfidence (Fischhoff, 1982a) found that none of the artifactual manipulations in Table 2 had produced consistent (or even any demonstrable) improvements in performance. It is an empirical question whether comparable reviews of the evidence for other biases (or additional studies of these two) will yield similar patterns of results. However, this review did offer some reassurance that apparent biases could not be accounted for (nor dismissed) quite so routinely. Until further evidence is in, the jury must remain out on these questions, with existing evidence providing

some guidance for the judgment that must come in its stead.

"It's True But You Shouldn't Say It"

Ascriptions of bias seem to be saying something fundamental--and rather unflattering--about people, namely, that they are not very good at consequential tasks. By implication, they also say something rather nice about the investigators in this domain. Not only do they know the right answers to questions where others stumble, but they are clever enough to design experiments illuminating those failures. For some observers, this is not only a dangerous combination, but one with ramifications far beyond the tidy little world of psychologists and their internecine struggles.

Quirks, in general, make good reading, meaning that they have a reasonable chance of being understood by the general public. These particular quirks regarding judgmental capabilities touch such an essential aspect of human affairs that it is likely to attract an audience having not only an interest, but a vested interest in their implications. For example, the role afforded the public in managing hazardous technologies should depend, at least in part, on their ability to understand the complex and uncertain technical issues involved. As a result, there are plenty of involved observers for whom "scientific proof" of lay people's foibles would constitute either very good or very bad political news, depending upon whether they wish to promote programs such as right-to-know laws and deregulation, which require public understanding, or consumer protection laws and deference to technical elites, which come in lieu of broad-based understanding (Fischhoff, Slovic and

Lichtenstein, 1982).

More generally, any sweeping statements about human capabilities reflect on how we think about our fellows or ourselves (if we take the research personally). There is a sharp distinction between the images expressed in Shakespeare's:

> "What a piece of work is man.
> How noble in reason, how infinite in
> faculties, in form and moving how
> express and admirable, in action
> how like an angel. In apprehension how like
> a god. The beauty of the world,
> the paragon of animals."

and Herbert Simon's:

> "The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world--or even for a reasonable approximation to such objective rationality [Simon, 1957; p. 198]."

one of the first comprehensive statements of concern about judgmental limitations (as juxtaposed by Slovic, 1972). Even with a very modest appraisal of psychology's influence on external events, one can still be concerned about its contribution to a climate of opinion that values people increasingly less. Or, one might just be offended by the effrontery of those academics who dare to make such pretentious statements.

A ready retort is that the scientists' responsibility is to tell it like it is, giving the most accurate account possible of what the data seem to say, integrating them with extant theory and surrounding them with appropriate qualifications. On intellectual grounds, doing anything else voluntarily implies violating the objectivity of science;

doing anything else involuntarily implies violating the independence of science (i.e., acceding to censorship). On practical grounds, as much of a disservice is done to people by giving them too much credit for decision-making capabilities as by giving them too little credit. In extreme forms, the former denies that they need help, whereas the latter denies that they can be helped.

Accurate transmission of the results of research carries, however, no assurance of accurate reception. Listeners' desire to hear a particular story about human judgment may heighten the normal processes by which scientific results are misunderstood under even noncontroversial conditions. Given the stakes riding on general statements about judgment, those attempting to make them have a special responsibility to hedge their conclusions properly. Moreover, they should monitor the way that those results are used, for cases where the hedges are either trimmed or magnified, either by those who fail to appreciate the niceties of experimental design or by those who choose to ignore them, in order to achieve some rhetorical purpose. Ensuring a balanced summary of results may also require retelling one's tale to different audiences, varying the presentation so as to avoid the different misunderstandings to which it is prone, a task which is not only tedious but seemingly at odds with scientists' usual mission of giving just the facts.

Such problems are inevitable whenever there is pressure to reduce a complex situation to sweeping statements. Contributing to listeners' confusion is the decided ambivalence of initial students of heuristics and biases. Their enterprise was part of the cognitive revolution in

psychology which showed, among other things, how suboptimal behavior that might be attributed to dark psychodynamic processes could also be accounted for by routine cognitive processes. This reinterpretation rescued people from one set of charges, at the price of calling into question another set of their faculties. The heuristics perspective also provided a fairly phenomenological account of behavior, in keeping with the cognitive revolution's disentanglement from behaviorism. Doing so ennobles people, in the sense of freeing them from the strictures of their environment, at the price of denying credit them for the (somewhat rudimentary) optimization that tight environmental control can assure.

Moreover, the basic heuristics-and-biases story is inherently convoluted, positing somewhat useful rules that can lead to predictable biases in well-characterized circumstances of indeterminate frequency. The listener wanting a simple message has grounds for seeing a glass half empty or a glass half full. Because it is hard to say anything general about the prevalence of conditions in which heuristics produce bias, the central studies in the area are mute regarding how "good" or "bad" people's judgment is. However, given the salience of quirks, it is easy to see how the biases could become a figure that dwarfs the ground created by the heuristics in routine (and reasonably successful) operation.

"It's True But It Doesn't Matter"

Some critics have been less timid in making general statements about what life is like, in order to be able to assess the impact that biases have on it. Sometimes conceding the legitimacy of the laboratory results, sometimes prefacing their work with "even if they are true,"

these critics argue that the biases are but epiphenomena, having little effect on people's lives. Although they might argue that the environment is benign, in the sense of not deliberately creating situations that fool people, a more common claim is that the environment is forgiving, in the sense of not penalizing people severely for those errors that they do make.

The most sophisticated versions of these claims create error theories, showing the sensitivity of particular classes of decision to particular kinds of error. For example, in an analysis that originally focused on the effect of perceptual imperfections on motor behavior, von Winterfeldt and Edwards (1982) showed that the expected value of many decisions with continuous options (e.g., invest X dollars) remains about the same even though people's estimates of parameters in their model of the decision problem vary moderately. Thus, neither modest errors, nor modest increments in information, will demonstrably affect what they get out of the situation.

Like any theoretical analysis, the power of that derived by von Winterfeldt and Edwards depends on the validity of its assumptions and the breadth of its applicability. Thus, additional analyses are needed to assess the impact of the biases wherever there are discrete decision options (e.g., operate/don't operate), wherever the magnitude of bias is beyond the (fairly broad) range specified here, wherever multiple instances of the same modest bias compound within a particular decision, wherever the bias affects aspects of decision making other than parameter estimation (e.g., problem structuring; von Winterfeldt & Edwards, 1975), or wherever people choose their options according to a

decision rule other than the expectation rule specified by this model.

For error theories to provide a generic strategy for restricting the importance of biases, credible accounts are needed for the variety of significant decision-making situations that people face, making the case that each kind of decision is insensitive to each kind of bias. To the extent that the range of potential problems can be severely constrained, then the general claim is strengthened. However, insensitivity itself carries a price. If biases do not affect most decisions, then there will be little sharp feedback regarding their existence, meaning that there will be little opportunity to learn to overcome them. As a result, one may be particularly vulnerable to their effects in those, perhaps few, situations in which they do matter.

As they emerge, such analyses provide important contributions to understanding the relationship between people and their environment. Their rigor is in sharp contrast to vague claims to the effect that experience would have taught people to avoid behavior that is frequently bad for them. Although a legitimate possibility, such claims are empty without specifying the mechanism of action whereby such learning takes place. One ready counterclaim is that life fails to provide the conditions needed for acquiring judgment as a learned skill. Table 3 asks about the possibilities that even experts have for learning to go beyond their data and rely on judgment. These conditions are met by U.S. weather forecasters making probability-of-precipitation forecasts (Murphy & Winkler, in press). At little or no personal cost, they receive prompt, unambiguous feedback from an institution that rewards them for facing the limits of their knowledge candidly. The resulting

forecasts are remarkably good (e.g., it rains 70% of the time they are 70% confident that it will rain). Whether other environments are similarly cooperative for particular judgments is a question for environment-by-judgment analysis. How cooperative they need to be is a question for empirical study.

Insert Table 3 about here

A discouraging feature of many biases as far as possibility of improving judgment is that, if real, they carry the seeds of their own propagation by disrupting the learning process. For example, overconfidence seems both to reflect and to foster relative disinterest in the reasons why a belief might be wrong; once undue interest in supporting evidence has shaped one's perceptions, there should be less perceived need to double check one's beliefs. Where such things happen, the chances diminish for activating another mechanism that could minimize the impact of biases: asking for help. Most people's lives seem to provide many more prompts to reach for a calculator or a dictionary than for a decision aid (or aide). Whether that is, indeed, the case is again an empirical question which can be answered, in part, in the lab and, in part, in the marketplace (by viewing the sales of different kinds of help).

The marketplaces for goods and ideas could, in principle, offer further avenues for muting the effects of bias, even without awareness of the need for help. For the marketplace for ideas to work, there would have to be a variety of views exchanged and a tendency for less biased ones to be more persuasive, rather than having the concurrence of

faulty beliefs enhance confidence in their rectitude and swamp wiser alternatives. For the marketplace for goods to work, one would have to describe processes by which competitive pressures cancelled, rather than magnified, the impact of biases, as well as prevented the exploitation of judgmental foibles by those with (even) less concern for people's welfare than the psychologists who "show them up" in the laboratory. At the moment, such analyses are lacking. Perhaps a better hope for biases cancelling one another lies in the ambiguity of how the heuristics are to be applied in particular situations, a topic considered below.

"You (The Experimenter) Have Got It All Wrong--People Are Doing Something Quite Different from What You Think and Are Doing It Quite Well."

Describing behavior as suboptimal presumes knowledge of what is optimal. Within decision theory, optimality is defined in terms of the quality of one's choices in the light of what one knows and wants. From this perspective, it would be wrong to claim suboptimality, for example, when people act overconfidently in a world that rewards overconfidence, or when they fail to do as well as we might have done given the benefit of our additional knowledge. When decisions are made in an unpredictable domain, one cannot judge the optimality of a decision by the desirability of the outcome that follows. Bad outcomes may happen to good decisions, and bad decisions may be followed, fortuitously, by good outcomes.

Table 4 summarizes these and additional factors that have frustrated attempts to evaluate the formal decision-making methods used in corporate and government arenas. All have analogs in evaluating the

more intuitive decision processes of individuals. Consider item 1, for example: the availability heuristic should not be unduly faulted for the poor results that it produces when used too casually (e.g., when an estimate is based on the first few instances or images that come to mind)--unless one feels that casualness is part and parcel of using heuristics (much as one might lose patience with a clinical treatment program whose failures were regularly explained away by the claim that practitioners had not applied it correctly).[3]

Insert Table 4 about here

Critics of the heuristics research have cited most of the factors in Table 1 as confounding investigators' assessments of the optimality of subjects' responses. Perhaps the most difficult charges to address are those associated with item 4, to wit: what constitutes a good outcome for subjects is not necessarily what would constitute a good outcome for the investigator, acting in their stead. The second of the six sections in Table 2 elaborates on this possibility, listing ways in which subject or experimenter might misunderstand one another's perspective on the task. That is, one might argue (a) that subjects deliberately sacrificed optimality in the task (as the investigator saw it) in order to achieve some alternative goal (e.g., learning about the system by making diagnostic mistakes), (b) that subjects share definitions of key terms (e.g., probability) different from those held or presumed by the investigator, (c) that the task was incompletely specified until subjects made additional assumptions that would have had to concur fortuitously with those made by the experimenter, or (d) that subjects

made a reasonable distinction (e.g., with regard to the relevance of data) to which the experimenter was insensitive.

It is to avoid such problems that the study of decision making has retreated to the laboratory where, in principle, all of details can be specified and all extraneous beliefs or distractions can be eliminated. Thus, one can tell subjects "we really want to see how well you can assess the extent of your own knowledge, not how confident an impression you can make on the strength of that knowledge." Yet, however careful investigators may be in designing their experiments to meet these requirements, no one conducts manipulation checks to ensure that every feature is understood exactly as intended--in part because of confidence in their ability to design experiments and divine subjects' interpretations, in part because of the difficulty of getting at how subjects construe a task at the moment it is performed, and in part because of the wilting number of possible construals that could be tested.

Were an investgator to run all these checks, they would still be restricted to subjects' construal of the elements appearing in that investigator's model of optimal behavior, that is, those features that need to be specified in order to determine what constitutes the optimal decision. Thus, those working within the Bayesian tradition (Fischhoff & Beyth-Marom, 1983) would have no natural interest in subjects' perceptions of the possibility for dynamic restructuring (whereby new evidence alters the set of hypotheses being evaluated) or in subjects' second-order probabilities (i.e., their level of confidence in their level of confidence), two features that are absent from the model (Diaconis & Zabel, 1982; Gärdenfors & Sahlin, 1982). They may also be

less than eager to check on construals whose incorporation in the model is cumbersome, such as interdependence in the interpretation of data, or the transaction costs and benefits associated with the decision-making process itself.

Thus, there is a lot of room for maneuver for those who would claim that subjects were really up to something other than what the experimenters intended. The question then becomes how much credibility to give these claims of optimality, relative to the claims of suboptimality that they are meant to counter. At the extreme, such counterclaims could represent no more than opportunistic special pleading, using whatever ad hoc reinterpretations are needed to make a point, without concern for the consistency of the claims made in different contexts or to their testability. In their dogged devotion to rationality as both a descriptive and prescriptive theory, economists lay themselves open to this trap. However, the analytical power that they derive from being able to use rationality in both ways means that economists have much more to lose by conceding some suboptimality than do psychologists, with their more descriptive enterprise. Indeed, one might see a case of throwing out the baby for the bathwater in the work of those theoreticians who have attempted to supplant time-honored models of rationality in order to account for a few apparently discrepant psychological results (e.g., Bell, 1985; Cohen, 1981).

There are three levels at which one could try to move counterclaims regarding what people are really doing in experiments from the realm of the possible to that of the plausible. One is the substantiation of claims regarding individual studies, through the detailed reading of

their protocols, secondary analysis of their results, and collection of supporting data. The second is to demonstrate the existence and importance of the psychological processes that have purportedly been ignored or excluded. For example, if one wished to argue that people may err deliberately in order to test their environment, then one might show that people can learn something from such probabilistic evidence, rather than being misled by its feedback (e.g., Einhorn, 1978; Einhorn & Hogarth, 1978). Finally, one can offer an alternative account for the entire existing body of data. Here, counterclaims are at a disadvantage, insofar as theories typically deal best with data whose creation they prompted. However, even there, these theories' account is never complete. They reign only so long as they do at least as well as their nearest competitor, meaning that there is room for imperfect alternatives (Lakatos, 1970).

"But Look At How Well People Do Other Things"

The depth and perversity of these biases seems strikingly at odds with the observation that, nonetheless, the world turns, planes fly, meals get on the table, and games of enormous complexity (e.g., chess, handball) are played. Could people be so good in one domain of complex behavior, yet so bad in the other? One answer to this question is that they couldn't, meaning that one of these seemingly conflicting observations is incorrect. Because the successes are an obvious fact of life, it then becomes natural to doubt the reality of the problems. They could be dismissed outright as laboratory curiosities (or fictions). Or, one could systematically increase the credibility of the various counterclaims discussed above, in order to chip away at the evidence of

apparent bias.

A more involved response is to step up the search for hither-to-neglected signs of prowess within the judgmental area itself, such as more valid entries in people's repertoire of heuristics (Nisbett, Krantz, Jepsen & Kunda, 1984) or ways in which a modest restructuring of judgmental tasks leads to better performance. The fifth section of Table 2 offers some generic forms of restructuring: (a) forcing respondents to express what they know explicitly, rather than letting it remain "in the head;" (b) encouraging respondents to search for discrepant evidence, rather than collecting details corroborating a preferred answer; (c) offering ways to decompose an overwhelming problem into more tractable and familiar components; (d) suggesting that respondents consider the set of possible situations that they might have encountered in order to understand better the specific situation at hand; and (e) proposing alternative formulations of the presented problem (e.g., using different terms, concretizing, offering analogies.)

Such restructuring presumes no guilt on the part of task or judge, but rather that there has been a mismatch, such that people have not brought their best cognitive skills to the task. These manipulations could be used as ways to help people overcome the limitations revealed in experiments (which would mean acknowledging their reality). Alternatively, one could argue that life itself tends to present problems in a (restructured) way that helps people to use their minds to best advantage.

As with other claims to the effect that life is more charitable to people than are experimenters, these need to be buttressed by direct

evidence or some convincing theory. One part of the evidentiary package is demonstrations that such restructuring works. One challenge to such demonstrations is posed by the very robustness of the biases they are meant to undo. In many such cases, any intervention that destabilizes the dominant response patterns will tend to reduce the degree of bias, making confusion look like wisdom. Although confusion is often an intermediate state between holding erroneous beliefs and holding appropriate ones, it can also be an intermediate state between two erroneous beliefs. Hence, some additional evidence is needed to show that a reduction in bias is accompanied by an increase in understanding.

Some years ago, we (Fischhoff, Slovic & Lichtenstein, 1979) believed that we had identified a simple form of restructuring which could be taught quickly and applied widely to judgmental problems. Noting that many biases represented the neglect of normatively important information (e.g., that concerning base rates, sample size), we thought the meaning of that information might be better understood if people considered how they would have responded had different values of it been reported. By and large, when subjects considered hypothetical alternative values before making an inference for the "actual" set of data, they showed much greater sensitivity to the commonly-neglected information than had subjects who received the simple value as one member of a single set of data. Although some subsidiary observations supported the interpretation that this manipulation had tapped latent understanding, we were disturbed that the improvement from it failed to generalize. In a subsequent study (Fischhoff & Bar-Hillel, 1984b), we found that the same manipulation was equally effective in sensitizing subjects to other

kinds of information that had been ignored appropriately in previous studies. Conceivably, the manipulation tapped some latent misunderstanding in these cases. However, a more parsimonious account is that it just shakes up the normal response patterns, which does little good unless accompanied by some substantive training in how to interpret data more adequately.

Facing the Biases

A final possibility is to abandon the search for a simple solution and look for an account that incorporates the kernels of truth in the original claims regarding biases, as well as the counterclaims and counter-counterclaims that they have generated. Acknowledging the coexistence of superlative performance in some areas, alongside deficits in others, opens a number of productive avenues. The first of these is reconciling the disparity. Just why is it that people are so good at some things and so bad at others? The set of possible reasons can be partitioned (along the lines of Table 2) into factors involving the tasks (and the environment in which they must be performed), factors involving the psychological capabilities that people bring to them, and factors involving the interaction between people and tasks.

The gist of the ambient account among researchers in the heuristics-and-biases tradition is that there are no fundamental psychological limitations, beyond information-processing capacity, that prevent people from acquiring the skills needed to make better judgments. Nonetheless, there are powerful environmental barriers to acquisition. One is the absence of systematic training in judgment or decision making in most jurisdictions (Beyth-Marom, Dekel, Gombo & Shaked, 1985). This defici-

ency would not be so serious were people's natural environment more forthcoming with opportunities for learning. However, as suggested above in the context of experts (Table 3), people seldom receive the orderly feedback needed to improve their judgment and decision making. In this regard, those are skills quite different than, for example, problem solving, where right and wrong answers are readily identified. Finally, there are barriers posed by the interaction between psychological processes and the environment. Fallible heuristics may survive because they generally have enough ecological validity to avoid the sort of egregious errors needed to get people even to think seriously about changing their ways (especially when they do not know where to look for help). Fallible heuristics may generate self-fulfilling prophecies when they prompt actions that serve to justify their predictions (Einhorn & Hogarth, 1978). And, they may be rewarded socially when they represent widely accepted patterns of judgment. From this general perspective, general statements about the prevalence and potency of bias are meaningless. However, there are strong pressures ensuring that they will arise (predictably, in certain situations) and that when they do their effects can be serious.

Although fairly controversial, this account of the place of bias in people's lives is still quite general, and certainly not the only one possible. Refining (or replacing) it will require a detailed, and potentially productive, analysis of the ecology of inference, considering the structure of tasks, the efficacy of heuristics, and the opportunities for learning. Given the precision and richness of the normative models for judgment and decision making, this area may offer parti-

cular promise for insightful theorizing. The error theories cited above are steps in this direction.

Taking the biases seriously, where they appear, is also a precondition for doing something serious about alleviating them. The middle sections of Table 2 describe ways of addressing the possibility that people have real problems, in this circumscribed domain. Optimistically, one could train them. The table describes an escalation design going through: mere warning about the possibility of bias without specifying its nature,[4] describing the direction (and perhaps extent) of the bias that is typically observed, providing a dose of feedback, personalizing the implications of the warning, and offering an extended program of training with feedback, coaching, and whatever else it takes to afford the respondent cognitive mastery of the task. The restructuring procedures of section 5 could be part of "whatever else it takes." Empirical evidence suggests that training does work, particularly when it is informed by the sort of psychological understanding about how people think (and how they might be helped to think better) that restructuring implies (Fischhoff, 1982a).

Where optimism is not rewarded, or the steps needed to achieve improvement cannot be afforded, it may be advisable to give up on people. One possibility is replacing people with some superior procedure, such as the use of regression models of clinical judgment (e.g., Dawes, 1979). A second is recalibrating fallible judgments to more appropriate values, assuming that the amount and direction of errors are predictable. For example, the (UK) Central Electricity Generating Board was better able to forecast power availability when it doubled its

engineers' chronically understated estimates of the time needed to return units to production (Kidd, 1970). A third possibility is acknowledging the imprecision in people's judgments when planning actions based on them, by being ready for surprises and open to alternative views.

Taking heuristics and biases seriously allows and forces one to expect more from them, such as precise predictions regarding when and how heuristics will be used. Many of the original demonstrations were tours de force, in which few observers would deny, say, that availability was the heuristic of choice and that it would be employed in a particular way. In other cases, though, several different heuristics might be used, each in several different ways (e.g., Fischhoff, 1975, pp. 297-8). For example, the relative availability of examples from competing categories might be quite different after a short search than after a long one, producing quite different estimates (Beyth-Marom & Fischhoff, 1977). With specification left so vague, the possibilities for interpretative ad hoc-ery could eventually rival those that frustrate attempts to test rationality as a descriptive model of human decision making. Short of, or as a step toward, a full-blown theory of heuristics application, one could develop procedures for assessing their usage in particular settings, as a basis for generating testable predictions for the summary judgments that should emerge (Fischhoff, 1983; Fischhoff & Bar-Hillel, 1984a).

Finally, taking the biases seriously should prompt serious questions regarding the nature of a full-fledged theory of heuristics. Currently, "heuristics and biases" is a broad umbrella for a diverse set of procedures whose common thread is the domain of tasks that evoke

them, rather than some fundamental underlying cognitive processes. The future could be the integration of these pieces into a comprehensive theory describing the allocation and direction of heuristics in different circumstances. Or, it could be the development of local theories, describing the cognitive processes evoked when there is recourse to heuristics. An adjunct to either kind of heuristics theory would be an account of what happens when people reject even these somewhat rudimentary procedures for dealing with uncertainty analytically. That is, what happens when people choose not to think their way through to a solution, drawing inferences through synthesis of available facts--but, rather, try to match new situations to old ones in order to find some habit, tradition, or other rule of thumb that will guide their behavior.

Such a mélange of principles and procedures at different levels may not be the kind of theory of judgment that we want. However, it may be the kind of account that is needed to treat the diverse ways that people cope with uncertainty.

Footnotes

1. In order to focus attention on issues, rather than personalities, these themes are discussed in generic terms rather than as positions offered by particular critics. Doing so also allows me to present a set of integrated alternatives, some of which have yet to be advanced, and to avoid attributing particular themes (or hybrids) to individual authors. Some points of access to the critical literature are Berkeley and Humphreys (1982), Cohen (1981, and commentary), Edwards (1982), Fischhoff (1982a,b; 1983), Hogarth (1981), Jungermann (1984), Kahneman and Tversky (1982a), Nisbett & Ross (1980), Phillips (1982).

2. For example, the theoretically proper procedure for motivating people to express their confidence in their own knowledge is the use of proper scoring rules (see Lichtenstein, Fischhoff & Phillips, 1982, for an elementary exposition and references to more technical ones). These take the form of algebraic rules that provide users with the highest expected value if they reveal their true beliefs. Unfortunately, it is not clear that people always accept the appeal of the expected value rule or can execute it if they try in such novel circumstances (Fischhoff, Goitein, & Shapira, 1982). What they do seem to learn from trial-and-error experience with the rule is that the rewards for expressing great confidence, while there, are pretty meager. As a result, only those determined to squeeze the last expected point out of the scoring rule will ever do so. By reducing confidence, the rule also reduces overconfidence, the most common result in such experiments. However, it is hard to argue that the sharpened incentives have produced sharpened self-understanding, rather than an artifactual adjustment.

3. Indeed, Table 4 was derived from considering the analogies between decision aiding and that other helping profession, psychotherapy (Fischhoff, 1980). In general, the standards of evidence demanded by clinicians seems to stack up quite well against those of the consultants selling decision-aiding services.

4. This strategy differs from inspiring people to work harder by implying that the potential error is systematic and that respondents need instruction, not just a fair chance.

Acknowledgements

This research was support by the National Science Foundation under Grant SES 8213452 to Perceptronics, Inc. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. Portions of this article were presented as colloquia to the Psychology Departments of the University of Wisconsin-Madison and Dartmouth College. The thoughtful comments of those audiences are gratefully acknowledged, as are those of Ward Edwards, Lita Furby, Robin Gregory, Robin Hogarth, Patrick Humphreys, Larry Phillips, Nils-Eric Sahlin, and Paul Slovic, for either making me think about these issues or helping me to think about them. Correspondence may be addressed to the author at Decision Research, 1201 Oak Street, Eugene, OR 97401.

References

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. Acta Psychologica, 44, 211-233.

Bell, D. (1985). Disappointment in decision making under uncertainty. Operations Research, 33, 1-27.

Berkeley, D., & Humphreys, P. C. (1982). Structuring decision problems and the 'bias heuristic.' Acta Psychologica, 50, 201-252.

Beyth-Marom, R., Dekel, S., Gombo, R., & Shaked, M. (1985). An elementary approach to thinking under uncertainty. Hillsdale, NJ: Erlbaum.

Beyth-Marom R., & Fischhoff, B. (1977). Direct measures of availability and frequency judgments. Bulletin of the Pschonomic Society, 9, 236-238.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? The Behavioral and Brain Sciences, 4, 317-331.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. American Psychologist, 34, 571-582.

Diaconis P., & Zabel, S. L. (1982). Updating subjective probability. Journal of the American Statistical Asociation, 34, 571-582.

Edwards, W. (1982). How to make good decisions. Acta Psychologica, 56, 7-15.

Einhorn, H. J. (1978). Decision errors and fallible judgment: Implications for social policy. In K. R. Hammond (Ed.), Judgment and decision in public policy formulation. Boulder, CO: Westview.

Einhorn, H. J., & Hogarth, R., M. (1978). Confidence in judgment: Persistence in the illusion of validity. Psychological Review, 85, 395-416.

Fischhoff, B. (1975). Hindsight≠Foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288-299.

Fischhoff, B. (1980). Decision analysis: Clinical art or clinical science? Operations Research, 28, 28-43.

Fischhoff, B. (1982a). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.

Fischhoff, B. (1982b). Latitudes and platitudes: How much credit do people deserve? In G. Ungson & D. Braunstein (Eds.), New directions in decision making. New York: Kent.

Fischhoff, B. (1983). Predicting frames. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 103-116.

Fischhoff, B., & Bar-Hillel, M. (1984a). Diagnosticity and the base-rate effect. Memory and cognition, 12(4), 402-410.

Fischhoff, B., & Bar-Hillel, M. (1984b). Focusing techniques: A shortcut to improving probability judgments? Organizational Behavior and Human Peformance, 34, 175-194.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. Psychological Review, 90, 239-260.

Fischhoff, B., Goitein, B., & Shapira, Z. (1981). SEU: A model of decision making. Journal of the American Society of Information Sciences, 32, 391-413.

Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S., & Keeney, R. (1981). Acceptable Risk. New York: Cambridge University Press.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance, 3, 552-564.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1979). Subjective sensitivity analysis. Organizational Behavior and Human Performance, 23, 339-359.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1982). Lay foibles and expert fables in judgments about risk. American Statistician, 36, 240-255.

Gärdenfors, P., & Sahlin, N. E. (1982). Unreliable probabilities, risk taking, and decision making. Synthese, 53, 361-386.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin, 90, 191-217.

Jungermann, H. (1984). The two camps on rationality. In R. W. Scholz (Ed.), Decision making under uncertainty. Amsterdam: North Holland (68-86).

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, 430-454.

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

Kahneman, D., & Tversky, A. (1982b). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases (pp. 201-208). New York: Cambridge.

Kidd, J. B. (1970). The utilization of subjective probabilities in production planning. Acta Psychologica, 34, 338-347.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.

Lakatos, I. (1970). Falsification and scientific research programs. In I. Lakatos & A. Musgrave (Eds.), Criticism and the growth of scientific knowledge. New York: Cambridge University Press.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: State of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press.

Murphy, A. & Winkler, R. (in press). Probability of precipitation forecast. Journal of the American Statistical Association.

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

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

Phillips, L. D. (1982). A theory of requisite decision models. Acta Psychologica, 56, 29-48.

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

Slovic, P. (1972). From Shakespeare to Simon: Speculation--and some evidence--about man's ability to process information. ORI Research Monograph, 12(2).

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. Journal of Experimental Psychology: Human Perception and Performance, 3, 544-551.

Taylor, S. E. (1982). The availability bias in social perception and interaction. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press, 190-200.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 4, 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

von Winterfeldt, D. (1975). Error in decision analysis: How to create the possibility of large losses by using dominated strategies (SSRI Research Report 75-4). University of Southern California, Social Science Research Institute, Los Angeles, CA.

von Winterfeldt, D., & Edwards, W. (1982). Costs and payoffs in perceptual research Psychological Bulletin, 91, 609-622.

Table 1

Heuristics and Biases

Heuristic	Bias	References
Availability: estimate the frequency of events by the ease with which instances are remembered; assess the likelihood by the ease with which it is imagined.	overestimation of unduly memorable or salient events	Tversky & Kahneman (1973), Kahneman & Tversky (1982b), Taylor (1982).
Representativeness: assess the probability of an event by how well it represents salient features of the data-generating process.	neglect of base-rate information; neglect of validity information.	Kahneman & Tversky (1972), Bar-Hillel (1980), Fischhoff & Bar-Hillel (1984a).
Anchoring: estimate numerical quantities by making a best guess and then adjusting in response to additional considerations.	undue influence of anchor; vulnerability to irrelevant anchors.	Tversky & Kahneman (1974).
Hindsight: make sense of outcome reports, integrating them with previous knowledge about event.	inability to reconstruct previous state of mind; unduly harsh judgments of decisions made in foresight.	Fischhoff (1975, 1982a), Slovic & Fischhoff (1977).
Belief confirmation: look for reasons supporting one's beliefs.	exaggerating extent of own knowledge.	Fischhoff, Slovic & Lichtenstein (1977), Koriat, Lichtenstein & Fischhoff (1980).

Table 2

Debiasing According to Underlying Assumption Methods

Assumption	Strategies
Faulty tasks	
Unfair tasks	Raise stakes Clarify instructions/stimuli Discourage second-guessing Use better response modes Ask fewer questions
Misunderstood tasks	Demonstrate alternative goal Demonstrate semantic disagreement Demonstrate impossibility of task Demonstrate overlooked distinction
Faulty judges	
Perfectible individuals	Warn of problem Describe problem Provide personalized feedback Train extensively
Incorrigible individuals	Replace them Recalibrate their responses Plan on error
Mismatch between judges and task	
Restructuring	Make knowledge explicit Search for discrepant information Decompose problem Consider alternative situations Offer alternative formulations
Education	Rely on substantive experts Educate from childhood

Source: Fischhoff (1982a).

Table 3

Conditions for Improving judgment

Do Experts Have Them?

1. Abundant practice with a set of reasonably homogeneous tasks. Experts should have such experience. They may use it to hone their judgmental skills or they may develop situation-specific habitual solutions, freeing themselves from the need to analyze (and think).

2. Clear-cut criterion events. Although experts are often required to make their judgments quite explicit, the objects of those judgments are often components of such complex (natural, social, or biological) systems that it is hard to evaluate the judge's level of understanding. Off-target judgments may be due to unanticipated contingencies, whereas on-target judgments may have been right for the wrong reason.

3. Task-specific reinforcement. Experts are, in principle, paid for performance. However, even when the wisdom of their judgments can be discerned, they may be rewarded on other grounds (e.g., did they bring good news? did they disrupt plans? did things turn out for the best?).

4. Explicit admission of the need for learning. Entering an apprenticeship program that confers expertise is surely a sign of modesty. Nonetheless, at every stage of that process and the professional life that follows it, certain advantages accrue to those who put on a good show and exude competence.

Table 4. Effects that Complicate Attempts to Evaluate the Efficacy of a Decision-Making Method

1. The fact that practitioners have been trained in a method and claim to be carrying it out is no guarantee that they are. Assessing fidelity of implementation is crucial for knowing what is being evaluated.

2. A well-designed method may fail because of unanticipated and uncontrollable changes in the world. Thus "good method" does not necessarily imply "good outcome."

3. Sometimes decision-making methods look good because they were fortunate enough to be used at times when they could not fail. Almost everybody and every method made money in the stock market of the 1950s and early 1960s. Thus "good outcome" does not necessarily imply "good method."

4. In some cases, defining a "good outcome" is far from a trival matter - for example, when one must weight short-term and long-term well-being.

5. The apparent success of some methods may be attributable less to their substance than to the atmosphere they create. These "non-specific treatment effects" include reduced anxiety, increased self-confidence, and heightened attention to the problem.

6. Anecdotal evaluations are often skewed by tendencies to be influenced by professional folklore and to interpret random fluctuations as consistent patterns.

7. An evaluation can be distorted by looking only for the positive effects a method produces and ignoring possible negative effects - or by looking only for the negative effects.

Source: Fischhoff, Lichtenstein, Slovic, Derby & Keeney (1981).