

Debiasing

Richard P. Larrick

The mind has its illusions as the sense of sight; and in the same manner as feeling corrects the latter, reflection and calculation correct the former.

Pierre Simon, Marquis de Laplace

Rationality and Debiasing

That the mind has its illusions is not without dispute. Thirty years of decision research has used rational theories from economics, statistics, and logic to argue that descriptive behavior falls systematically short of normative ideals. But this apparent gap between the normative and the descriptive has provoked many debates: Is there in fact a gap? And, if there is, can it be closed – that is, can biases be “debiased”?

Many economists and philosophers have argued on principle that there is no gap: people are essentially rational, any errors are random and non-systematic, and apparent systematic discrepancies are attributable to improper empirical methods. Stanovich (1999) has aptly termed this group the Panglossians. Early research on debiasing largely served to counter the Panglossian position by demonstrating the robustness of systematic biases to various corrective measures (Fischhoff, 1982). The existence of systematic biases is now largely accepted by decision researchers, and, increasingly, by researchers in other disciplines.

Accepting the existence of a normative–descriptive gap raises the question of how the gap might be closed. One approach has focused on increasing the motivation to perform well. A critical assumption in this approach is that people possess normative strategies and will use them when the benefits exceed the costs. The remaining approaches do not presume this. Instead, they assume that intuitive strategies are imperfect, but that they can be replaced by strategies that approach normative standards (even if falling short).

The identification and dissemination of better strategies is known as prescriptive decision making.

My emphasis will be mainly on prescriptive strategies that individuals themselves can adopt, as opposed to techniques used by external agents to modify the decision environment. An example of the former approach would be increasing retirement savings by training people on the principle of compounding, either abstractly or as a rule of thumb (e.g., “the rule of 72” – an investment that grows at X percent per year will double roughly every $72/X$ years). An example of the latter approach is Thaler and Benartzi’s (2001) highly successful demonstration that organizations can rebias employees to save more by changing the status quo and by exploiting mental accounting. Both rebiasing (using one bias to offset another) and changing the decision environment are viable methods for debiasing. However, I will focus on equipping individuals with strategies because this approach tends to increase their decision skills and their ability to apply those skills to new decision domains (in this example, perhaps, college savings).

One approach to prescription focuses on modifying the *cognitive* strategies of the individual. In this view, optimal prescriptive strategies represent a compromise between a strategy that approximates the normative ideal, but that can be remembered and implemented given ordinary cognitive limitations on memory and computation. Successful implementation also requires an encoding strategy for recognizing when to apply a cognitive strategy (Nisbett, Krantz, Jepson, & Kunda, 1983).

The extent to which purely cognitive strategies can improve reasoning is a source of debate. Stanovich (1999) terms those optimistic about improving cognitive strategies Meliorists, who believe that everyday reasoning falls far short of the ideal, but can be improved through experience and education (Nisbett, 1993). By contrast, the Apologists perceive supposed normative standards to be unattainable for many tasks because of computational constraints and unnatural problem representations (such as the use of Bayes’ rule in probabilistic reasoning) and reject the supposed normative standard. They argue that the basic intuitive strategies people follow are evolutionarily well adapted to naturalistic judgment tasks (such as reasoning about frequencies) (see Gigerenzer, Chapter 4, this volume). Each view has different implications for prescription: Meliorists are optimistic about improving cognitive strategies through training, whereas Apologists suggest that decision tasks must be matched to evolutionarily-adapted strategies (I return to an intriguing intermediate position by Sedlmeier, 1999).

The degree to which reasoning can be improved through cognitive strategies has important implications for philosophical debates about rationality. Many philosophers are reluctant to equate rationality with strategies that are not humanly achievable, in which case, the most accurate cognitive strategies that people can use become the standard for rationality. As Stich wryly noted, “it seems simply perverse to judge that subjects are doing a bad job reasoning because they are not using a strategy that requires a brain the size of a blimp” (cited in Stanovich, 1999). Stanovich (1999, in work conducted with West) provides interesting evidence on this point in support of the Meliorists. They observe that, contrary to the Apologist view, there is always a subset of decision makers who give a normative response on a decision task, indicating that at least some people have it in their repertoire. Moreover, the pattern of normative responses across individuals is systematic. Normative responses are correlated positively with general aptitude (see

Jepson, Krantz, & Nisbett, 1983; Larrick, Nisbett, & Morgan, 1993) and with each other across highly diverse decision tasks. These data represent an existence proof that normative cognitive strategies are not unattainable, but are systematically held and used by some individuals.

Although identifying cognitive strategies that individuals are able to implement informs rationality debates, it is not the only means to close the normative–descriptive gap. A second approach to prescription is to expand possible strategies to include techniques external to the decision maker. This represents a Technologist alternative to the Meliorist–Apologist debate: individual reasoning can approach normative standards through the use of tools. Such techniques include using groups in place of individuals, improving information processing through decision aids and information displays, supplementing intuitive decision making with formal decision analysis, and replacing individual judgment entirely with statistical models. Debates about rationality have focused on purely cognitive strategies, obscuring the possibility that the ultimate standard of rationality might be the decision to make use of superior tools.

Despite the different emphases in these approaches, they share a common implication: debiasing requires intervention. Laplace optimistically observed that people recognize and correct their own biases, but there are many reasons to doubt that lone individuals can debias themselves (Kahneman, 2003). In part, this is a matter of which phenomena are declared biases – biases that are difficult for individuals to recognize and correct are selected into the canon of judgment and decision-making research, whereas those that are easily recognized and corrected are not. But there are other reasons why individuals are not able to debias themselves. First, they will often not realize when they have used a poor decision process – feedback on their decision outcome may be delayed, or the causal determinants of the outcome may be ambiguous, making both the existence and source of error difficult to identify (Hogarth, 2001). Second, the tendency to use decision outcomes to evaluate decision processes can lead to faulty conclusions in decisions made under uncertainty. These conclusions may be distorted further by self-serving attributions of ability that lead decision makers to attribute good outcomes to skill and poor outcomes to situational factors.

The study of debiasing, therefore, must go beyond identifying better strategies to identifying methods for *equipping* individual decision makers with those strategies. This is where the traditional study of individuals in isolation may underestimate the potential for improving decision making. Even though lone individuals do not debias themselves, they are surrounded by cultural mechanisms that compensate for their shortcomings (Camerer & Hogarth, 1999; Heath, Larrick, & Klayman, 1998). For example, the evolution of normative models over the last two centuries not only *revealed* intuitive shortcomings, but provided the disciplinary knowledge that could *repair* them (Nisbett et al., 1983). The result is an ongoing race between the identification of biases and the diffusion of tools for reducing them. There is no guarantee, of course, that standard economics and statistics curricula provide the best means for improving intuition (Hogarth, 2001). Part of the research agenda for debiasing, therefore, needs to be identifying methods that promote the acceptance and use of superior decision strategies, which I return to in the final section.

The Nature of Biases

One way of organizing a review of debiasing is by specific bias (Fischhoff, 1982), but the field's success at generating new biases makes this approach impractical. Fortunately, an exhaustive list of biases is not necessary. In a classic article on debiasing, Arkes (1991) argued that a few general causes underlie a wide range of biases, and that understanding these causes facilitates identifying when different debiasing strategies will be effective. His first two categories are errors that are attributable to unconscious, automatic System 1 processes (Kahneman, 2003; Kahneman & Frederick, 2002; Stanovich, 1999). The third category is errors attributable to more conscious and deliberative System 2 strategies (see Table 16.1):

- *Psychophysically-based error (System 1)*: This category includes errors produced by non-linear translations of stimuli in judgment and evaluation. The prototypical examples are reference point effects (see Wu, Zhang, & Gonzalez, Chapter 20, this volume), in which a reference point introduces a kink in slope (due to loss aversion) and curvature (due to diminishing sensitivity) in valuation. Because reference points can shift across contexts – depending on what comparisons are perceptually salient or accessible in memory – the same stimulus can be judged inconsistently (Kahneman, 2003).
- *Association-based error (System 1)*: This category includes errors that are caused by automatic processes that underlie the accessibility of information in memory (Kahneman, 2003). Association-based errors occur when an initial representation, often evoked by a stimulus, leads to the activation of conceptually or semantically associated cognitions and the inhibition of unassociated cognitions. A major consequence of association-based processes is the recruitment of a narrow and often biased information base from which to make judgments and decisions (Payne, Bettman & Schkade, 1999), including narrow framing (Kahneman, 2003).
- *Strategy-based error (System 2)*: The third category includes errors caused by the use of inferior strategies or decision rules. In my use of this category, I will depart from Arkes's (1991) original assumption that strategy-based errors are adaptive – that is, that they reflect a rational benefit–cost calculation. Although there is substantial evidence showing that people adapt their decision strategy to situational demands (Payne, Bettman, & Johnson, 1993), there is little direct evidence that they select strategies optimally or gauge effort and accuracy accurately (Fennema & Kleinmuntz, 1995). Moreover, people simply may not have the normative strategies in their intuitive repertoire, in which case reliance on inferior strategies is not a calculated choice, but a necessity. System 2, in this view, can itself be a major source of error if it contains either flawed strategies or poorly-calibrated strategies that produce under- or overcorrection (Wilson & Brekke, 1994).

Of course, many biases are multiply-determined (see the sunk cost examples in Table 16.1). The implication is that there is unlikely to be a one-to-one mapping of causes

Table 16.1 General causes of bias

<i>General causes of bias</i>			
	<i>Psychophysical</i>	<i>Associationistic</i>	<i>Strategy</i>
Description	Non-linear translation of stimuli in judgment and evaluation	Activation of conceptually or semantically associated cognitions	Inappropriate rules
Example biases	<ul style="list-style-type: none"> • Status quo bias • Preference reversals due to joint vs. separate evaluation • Curvature of the probability weighting function 	<ul style="list-style-type: none"> • Some forms of anchoring • Some forms of confirmation bias • Hindsight bias • Subconscious priming 	<ul style="list-style-type: none"> • Positive test strategies • Lexicographic choice rules
Shortcomings	<ul style="list-style-type: none"> • Non-linear translation of dimensions that are assumed linear in normative theories (e.g., probability). • Inconsistent judgment of a stimulus 	<ul style="list-style-type: none"> • "Functional fixedness" in problem solving • Overweighting focal outcomes in probability judgment • Narrow recruitment of attributes and alternatives in choice 	<ul style="list-style-type: none"> • Neglect of relevant information in judgment and choice • Improper combination of inputs in judgment and choice
System	System 1	System 1	System 2
Sunk cost example	Diminishing sensitivity (convexity) in losses makes additional losses less painful after initial sunk cost	Cognitions that are consistent with the initial investment decision are accessible, inflating $P(\text{success})$	Incorrect mental accounting rule: Past costs are kept in current accounts

to bias, or of bias to cure. Different processes may also be interdependent. Processes in the strategy category often rely on the output of the other processes, such as when a combination rule uses non-linear transformations of probabilities as inputs.

I will refer to these categories in discussing when different debiasing approaches are effective. The next sections present arguments and findings on major debiasing approaches. The strategies are organized according to the three approaches outlined in the introduction: Motivational, cognitive, and technological strategies. The final section discusses the major issues underlying the selection, diffusion, and implementation of different debiasing strategies in practice.

Motivational Strategies

Incentives

Economists have often responded to claims about decision errors with a call for better incentives. The assumption is that individuals will expend more effort on “reflection and calculation” – that is, System 2 will kick in – if the stakes are high enough. There is little empirical evidence, however, that incentives consistently improve mean decision performance (see Camerer & Hogarth, 1999, for a selective review). For example, early studies of incentives found that real stakes strengthened preference reversals (see Hsee, Zhang, & Junsong, Chapter 18, this volume). Subsequent studies of other biases found that incentives reduced biases in only a handful of cases. Camerer and Hogarth (1999, p. 33) reached the provocative conclusion that “there is no replicated study in which a theory of rational choice was rejected at low stakes in favor of a well-specified behavioral alternative, and accepted at high stakes.”

To understand why incentives are generally ineffective, a second assumption in the incentives approach needs to be recognized. For incentives to improve decision making, decision makers must possess effective strategies that they either fail to apply or apply with insufficient effort when incentives are absent. In the words of Camerer and Hogarth (1999), decision makers must possess the necessary “cognitive capital” to which they can apply additional effort. They note that incentives do improve performance in settings such as clerical and memorization tasks, where people possess the cognitive capital required to perform well but lack the intrinsic motivation. Few decision tasks, however, are analogous to simple clerical work or memorization. Instead, experimental decision-making tasks are either quite complex (e.g., requiring the use of Bayes’ rule, which few people intuitively possess); or they are relatively simple, but require that a decision maker possesses both the right strategy (such as the conjunction rule) *and* the ability to recognize when to apply it. When decision makers lack the necessary cognitive capital, incentives may lead them to apply inferior strategies with more determination, producing a pattern I will call the “lost pilot” effect (“I don’t know where I’m going, but I’m making good time”).

A few decision tasks do benefit from greater effort being applied to simple strategies. In multiattribute choice, accuracy incentives lead people to search more extensively for information and to process information more by alternative than by attribute, resulting in more accurate decisions (Stone & Ziebart, 1995; see also Creyer, Payne, & Bettman, 1990). However, responding to incentives by using more information and by changing strategies can produce a “lost pilot” effect on some tasks, especially *stochastic* tasks. For example, incentives lead decision makers on prediction tasks to rely less on base-rate information and more on imperfect cues that they use inconsistently as they “chase” error, often *reducing* performance on these tasks (Arkes, Dawes, & Christensen, 1986; Hogarth, Gibbs, McKenzie, & Marquis, 1991).

Arkes (1991) has argued that the automatic nature of association-based and psychophysically-based errors should make them largely unresponsive to incentives. This

has held true for most biases of these types (e.g., hindsight bias, overconfidence, framing effects). Surprisingly, however, incentives have been shown to reduce the influence of anchors in some instances (see Epley, Chapter 12, this volume). Social psychologists have proposed that people hold intuitive theories about some association-based biases and can recognize and deliberately adjust for them (Wegener & Petty, 1995; Wilson & Brekke, 1994). Thus, incentives can increase the effort decision makers expend in correcting association-based biases – *if* decision makers recognize when they occur (Stapel, Martin, & Schwarz, 1998). A promising area for future research is identifying when people *spontaneously* apply an intuitive theory of association-biased errors.

Although incentives have been ineffective at reducing most biases in laboratory studies, these results may not reflect the true potential of incentives outside the laboratory. First, lack of effort may be a serious problem in some organizational decisions where tasks truly are boring. If the individual has little at stake, he may be satisfied with a superficial search of alternatives, attributes, and cues. In this setting, incentives may be a useful tool to improve decision making and align individual effort with organizational interests. Second, although incentives cannot improve cognitive capital in the course of a brief experiment, they can motivate people to acquire the decision skills they need over longer periods of time (Camerer & Hogarth, 1999).

Accountability

A second motivational approach to debiasing is holding people accountable for their decisions – that is, giving them the expectation that they will later have to explain their decision to others. The logic of accountability is similar to the logic of incentives, except that it depends on the motivational effects of social benefits (such as making a favorable impression and avoiding embarrassment). The principal mechanism by which accountability improves decision making is pre-emptive self-criticism. In preparation for justifying their decisions to others, decision makers anticipate the flaws in their own arguments, thereby improving their decision processes and outcomes.

The popularity of the accountability paradigm has led to many tests of accountability effects on various biases, with notable successes (see Lerner & Tetlock, 1999, for an excellent review). As with monetary incentives, accountability primarily improves performance on tasks for which people already possess the appropriate strategy (Lerner & Tetlock, 1999), such as the sunk cost rule among MBA students (Simonson & Nye, 1992). Just as with monetary incentives, accountability leads to greater effort (e.g., time spent on a task) and use of information (e.g., information searched in an information display) (Huber & Seiser, 2001), which may often lead to improved performance. But, just as with monetary incentives, the use of more information leads to a “lost pilot” effect on prediction tasks if cues are unreliable (Siegel-Jacobs & Yates, 1996, Study 1; Tetlock & Boettger, 1989).

The social nature of accountability makes it different from monetary incentives in several ways. One interesting difference is that accountability evokes a strong social need to look consistent to others. Although rigid consistency is detrimental for many tasks, it does improve prediction tasks, where the inconsistent weighting of reliable cues is a

major source of error (Siegel-Jacobs & Yates, 1996). The social nature of accountability also introduces some potential problems. First, accountable decision makers tend to “give the people what they want.” If they know their audience’s preference for a specific decision *outcome*, decision makers distort their decision process to justify that outcome; if they know their audience’s preference for a decision *process*, they are more likely to use that process (Brown, 1999). Consequently, justifying a decision to an audience with *unknown* preferences leads to pre-emptive self-criticism, but justifying a decision to an audience with *known* preferences leads to biased rationale-construction. Second, the focus on justification may have the effect of exacerbating justification-based decision biases (see Shafir and Leboeuf, Chapter 17, this volume). For example, both attraction and compromise effects are amplified by accountability (Simonson, 1989). Accountability is likely to strengthen reliance on salient or easily justified dimensions, such as outcome probabilities in choice.

Cognitive Strategies

“Consider the opposite”

By necessity, cognitive strategies tend to be context-specific rules tailored to address a narrow set of biases, such as the law of large numbers or the sunk cost rule. This fact makes the simple but general strategy of “consider the opposite” all the more impressive, because it has been effective at reducing overconfidence, hindsight biases, and anchoring effects (see Arkes, 1991; Mussweiler, Strack, & Pfeiffer, 2000). The strategy consists of nothing more than asking oneself, “What are some reasons that my initial judgment might be wrong?” The strategy is effective because it directly counteracts the basic problem of association-based processes – an overly narrow sample of evidence – by expanding the sample and making it more representative. Similarly, prompting decision makers to consider alternative hypotheses has been shown to reduce confirmation biases in seeking and evaluating new information.

Soll and Klayman (2004) have offered an interesting variation on “consider the opposite.” Typically, subjective range estimates exhibit high overconfidence. Ranges for which people are 80 percent confident capture the truth 30 percent to 40 percent of the time. Soll and Klayman (2004) showed that having judges generate 10th and 90th percentile estimates in separate stages – which forces them to consider distinct reasons for low and high values – increased hit rates to nearly 60 percent by both widening and centering ranges.

“Consider the opposite” works because it directs attention to contrary evidence that would not otherwise be considered. By comparison, simply listing reasons typically does not improve decisions because decision makers tend to generate supportive reasons. Also, for some tasks, reason generation can disrupt decision-making accuracy if there is a poor match between the reasons that are easily articulated and the actual factors that determine an outcome (Wilson & Schooler, 1991). Lastly, asking someone to list too many contrary reasons can backfire – the difficulty of generating the tenth “con” can

convince a decision maker that her initial judgment must have been right after all (see Roese, Chapter 13, this volume).

Training in rules

An important issue in rationality debates is whether people's inferior strategies can be replaced by better strategies. A practical question then arises: *How* do you replace them? Experience is one possible method, but it is often an inexact and even misleading teacher (Hogarth, 2001). A second method, training, is potentially more precise.

In an extensive program of research, Nisbett (1993) and his colleagues explored the effectiveness of training on normative rules, leading to two sets of implications for debiasing. First, their research identified specific cognitive factors that facilitate the learning and use of normative rules (Fong & Nisbett, 1991; Fong, Krantz, & Nisbett, 1986; Jepson et al., 1983; Nisbett et al., 1983). Second, their research demonstrated that formal training in basic disciplines, such as economics and statistics, is an important cultural mechanism for transmitting effective cognitive strategies (Fong et al., 1986; Larrick, Morgan, & Nisbett, 1990; Lehman & Nisbett, 1990) – although the transmission process can no doubt be improved (Nisbett, 1993).

A basic assumption underlying this work was that people often have a rudimentary understanding of statistical, logical, and economic principles, but have difficulty in knowing how and when to apply them. For example, Nisbett et al. (1983) argued that people have an understanding of basic statistical principles, such as the tendency for sample means to reflect the population mean more accurately as samples get larger. They argued, however, that this understanding is better developed in transparently probabilistic domains than in other domains, but can be improved with experience. When sports novices were told about a new player who had a great performance during a team tryout but performed less well during the season, they often gave deterministic explanations, such as "once the player made the team, he slacked off"; by contrast, sports fans were more likely to attribute the pattern to the small, unreliable sample of evidence provided in the tryout. Fong et al. (1986) went on to demonstrate that principles of sampling and sample variability could be taught in short training sessions either abstractly or with concrete examples, and that a combination of the two was most effective. Finally, Fong and Nisbett (1991) demonstrated that decision makers trained in one type of domain (e.g., sports performance) successfully generalized the rule to other domains (e.g., test taking), although cross-domain transfer diminished over two weeks.

A series of other laboratory studies focused on training in logical and economic principles. In research on logical reasoning, Cheng, Holyoak, Nisbett, and Oliver (1986) successfully trained undergraduates to reason with the material conditional (*if p, then q*), where verification of the relationship requires that a decision maker examine evidence regarding *p* (to test whether *q* is true) and *not-q* (to test that *p* was *not* true) – a pattern of testing that is rarely observed in the Wason selection task. Training was more effective using familiar, pragmatic rules about permission and obligation (if you drink alcohol, you must be at least 18 years of age) than using the purely abstract rule. In addition, combining abstract principles with concrete examples proved particularly important for

learning this rule. In research on economic principles, Larrick et al. (1990) demonstrated that students could be trained to ignore sunk costs in financial domains and generalize the rule to time decisions (and vice versa), and that they could correctly distinguish between sunk cost problems for which the normative principle implies opposite actions (discontinuing versus continuing investments, respectively).

Finally, a series of cross-sectional and longitudinal studies demonstrated differential effects of *disciplinary* training. Economics professors were more likely than biology or humanities professors to report abandoning a consumer activity in which they had “sunk” money, such as watching a movie or eating a restaurant meal, despite having the same consumption opportunities (Larrick et al., 1990, 1993). And students majoring in social science and psychology show improved reasoning on statistical problems after three years of related coursework, but no improvement on unrelated logical rules (Lehman & Nisbett, 1990).

Overall, these studies demonstrated that classes of decision rules could be taught effectively, often with relatively brief training. The most effective approaches combined an abstract principle with concrete examples, where experience with examples provided skills at mapping the principle to specific content. Examples-training is important because improvement is not just a matter of enhancing the strategies in System 2, but making their use automatic – in effect, making recognition of when to use them as a System 1 process. The rules that were taught successfully were either relatively simple, such as the sunk cost rule, or familiar, such as the law of large numbers. It is important to note that this rule-training research did not tackle highly complex, unfamiliar, abstract rules, such as Bayes’ rule; the assumptions underlying this research would suggest that Bayes’ rule would be a poor candidate for training.

Training in representations

A second program of training (Sedlmeier, 1999) was inspired by research showing that people reason more accurately about frequencies than about probabilities (Gigerenzer & Hoffrage, 1995). For example, Tversky and Kahneman (1983) showed that the conjunction fallacy occurs less when reasoning about a set of instances than when judging a single case (see Chapters 1 and 4, this volume; also Griffin & Buehler (1999) on why frequency formats are not a panacea). The relative effectiveness of reasoning about frequencies illustrates a general debiasing principle of “moderator as repair”: when a variable is found that moderates accuracy on a decision task, it can become the basis for designing a debiasing strategy. In this case, two strategies are possible. A technological strategy is to present information to decision makers as frequencies rather than as probabilities, thereby debiasing the environment. A cognitive strategy, pursued by Sedlmeier (1999), is to train people to translate probabilistic reasoning tasks into frequency formats.

Through a computer-based set of instructions and illustrations, Sedlmeier trained participants to use a probability-based *or* frequency-based approach to solve a *probability* problem, and then tested performance several weeks later. On the conjunction rule, both probability training (on Venn diagrams) and frequency training (on frequency grids)

proved highly effective. For problems that required reasoning about conditional probabilities or Bayes' rule, frequency training (on frequency grids and frequency trees) proved highly effective and durable, surpassing the effects of probability training. The most noteworthy part of this impressive research is that subsequent test problems were always presented in *probability* terms; thus, participants' success on later tests showed that they learned to apply the frequency tools to novel, single-case probability questions. Sedlmeier's training techniques have important implications for making statistics classes useful to everyday decisions.

Training in biases

Research on behavioral decision theory (BDT) is increasingly taught in psychology, law, and management curricula. "Stupid human tricks," as a friend has called them, are often taught in these classes to demonstrate inconsistencies in human reasoning, but with no accompanying instruction in how to overcome them, except a warning such as "beware availability." It would be interesting to perform a controlled experiment to test whether BDT courses reduce decision biases, as statistics courses increase the use of some statistical reasoning (Fong et al., 1986). Just as statistics and economics classes often miss the opportunity to develop people's intuitions through behavioral examples, courses that contain behavioral decision research may miss an opportunity to improve people's intuitions if they do nothing more than demonstrate the flaws. Without accompanying recognition skills and decision tools, it is unlikely that "awareness" alone would be sufficient.

Technological Strategies

Group decision making

Groups are often disparaged as decision-making resources because social influence processes undermine their effectiveness. People in groups often intentionally withhold or misrepresent their private judgments to avoid the social costs of rejection or to "free ride" on the efforts of others. But perhaps the most insidious problem in groups is that people are unknowingly influenced by the public judgments of others. Especially under conditions of uncertainty, people are susceptible to anchoring on the judgments of others in forming their own judgments.

Despite these problems, there are many reasons that groups might be beneficial. First, groups serve as an error-checking system during interaction. Second, "synergies" can emerge when people with complementary expertise interact. But the main benefits of a group may not depend on interaction at all. The third and arguably most important reason that groups improve decision making is statistical. Groups increase the effective sample size of experience used to make a decision. The result is that on tasks that require novel solutions – such as creativity or hypothesis generation tasks – groups hold more

diverse perspectives than any one individual. And on tasks that require estimation – such as forecasting or evaluation tasks – the larger sample and diversity of cue-usage in groups makes the combination of individual judgments a powerful way to reduce individual error. The result is that simply averaging individual forecasts has proven a robust method of reducing errors in prediction and estimation (Clemen, 1989; Soll & Larrick, 2004).

However, the statistical benefit of aggregation is fragile. The effective sample size of a group is greatly reduced to the extent that group members' judgments have shared errors (Hogarth, 1978) – in the extreme, each person becomes “redundant” with any other. Shared training, shared experiences, and shared discussions all lead group members to hold a similar view of the world – *and* similar blind spots. Although holding a similar view can foster group cohesion, it reduces the informational benefits of group decision making.

Using groups to improve decisions ultimately depends on assembling a group with diverse experiences and training, and then following a process that preserves the diversity of perspectives. If run effectively, groups generate their own “consider-the-opposite” process. It is interesting that the most popular group decision-making method – brainstorming – comes up wanting on preserving diversity. Brainstorming is designed to encourage diversity of ideas by separating a “no-criticism” idea-generation phase from a selection phase. Despite these helpful rules, the number of unique, high-quality ideas produced in an n -person brainstorming group does not come close to matching the output of n people working separately for the same period of time (dubbed “nominal groups”). The principal explanations for this deficit have been evaluation apprehension and the sequential rather than simultaneous pooling of information. But an uninvestigated flaw in brainstorming is that early suggestions tend to “contaminate” all members' subsequent ideas. The fundamental requirement of group decision making is that individuals must formulate their own hypotheses, judgments, and estimates *independently* of each other before working in a group; once into the group process, shared ideas can spark new insights.

Linear models, multiattribute utility analysis, and decision analysis

In 1772, Benjamin Franklin proposed to his friend Joseph Priestley a “moral or prudential algebra” for making difficult decisions. It entailed dividing a sheet of paper into two columns – Pro and Con – and then listing examples of each over the course of a few days because “all the reasons pro and con are not present to mind at the same time; but sometimes one set present themselves, and at other times another, the other being out of sight.” Once pros and cons are fully enumerated, sets can be compared; when a subset in one column has the same “respective weight” as a subset in the other, both can be struck out, until the decision maker can tell “where the balance lies” (from Dawes & Corrigan, 1974).

Franklin's proposal is celebrated as the forerunner of modern decision analysis because at its core is the basic tenet of decision analysis: “Decompose a complex problem into simpler problems, get one's thinking straight in these simpler problems, [and] paste these analyses together with logical glue . . .” (Raiffa, 1968). For example, a hiring

decision can be decomposed into *alternatives* (e.g., candidates under consideration), *attributes* (e.g., characteristics of the candidates, such as teaching experience), and *attribute levels* (e.g., such as three years of teaching experience). Each attribute needs to be weighted and then combined across each alternative. This commonly takes the linear, additive form of $U(A) = wx_1 + wx_2 + wx_3 + \dots + wx_n$, where w is the weight assigned to an attribute and x is the attribute level.

Proper and improper linear models

The simplest way to derive weights for a linear model is to use a statistical technique such as multiple regression analysis to fit a criterion to a set of predictor variables. In a series of influential reviews, Meehl and Dawes (see Dawes, Faust, & Meehl, 1989) demonstrated that, across scores of studies, statistical models based on past data consistently outperformed the “holistic” estimates of human judges on new cases (even when human judges had access to the same – or more – attribute information as the statistical model). Of course, a “proper” statistical model necessarily captures the true linear relationships in a set of data and therefore represents the upper bound that a perfectly linear, additive human judge could attain. More surprising is that a variety of “improper” models not based on past data also outperform intuitive judgment (Dawes & Corrigan, 1974). These include “bootstrapped” models (see Schoemaker, Chapter 14, this volume), in which a judge’s holistic judgments are regressed on a set of attributes, capturing the judge’s inevitably flawed attribute weights (or “policy”); and equal-weight models, in which a set of relevant attributes are identified, attribute values are normalized, and then given equal weight in an additive model (Dawes & Corrigan, 1974). Camerer (1981) has shown that bootstrapped models outperform holistic judgment consistently – if not dramatically – and that equal weights perform as well as or better than bootstrapped weights under a wide range of conditions (see also Payne et al., 1993).

Why do blatantly improper models outperform intuitive judgments? One reason is that, in Franklin’s terms, relevant attributes are often “out of sight.” In tasks that rely on recall, accessibility is likely to make some attributes more salient than others. But even with all relevant attributes available, attention is prone to wander, leading judges to focus on different attributes and to weight them differently as they evaluate specific alternatives, producing cognitive inconsistency (Hammond & Summers, 1972; see Goldstein, Chapter 3, this volume). People may also cope with an overabundance of attributes by using a non-compensatory strategy, such as elimination by aspects, to simplify the task (see Payne and Bettman, Chapter 6, this volume). Thus, even an *improper* linear model is effective because it ensures that *all* the attributes are used, and that they are *weighted* and *combined* consistently. A biased model consistently applied is an improvement over a biased and inconsistent human. Overall, linear models are ideally suited for tasks in which there are a large number of alternatives to review. It is precisely such data-rich but repetitive tasks that prove the most taxing on human information processing and benefit the most from substituting a model for a human.

Multiattribute utility (MAU) analysis

An alternative method for assigning attribute weights in a linear model is to elicit them directly from a decision maker through MAU analysis (see Pidgeon and Gregory,

Chapter 30, this volume). For unique choice problems that have no historical precedent (or accuracy criterion) this is the main alternative to linear models. A variety of methods exist for eliciting weights (Clemen, 1996), and the methods tend to yield similar estimates, indicating reliability (Leon, 1997; Stillwell, Barron, & Edwards, 1983). A practical implication of having multiple methods available is that they can be used simultaneously, and discrepancies examined and reconciled (Payne et al., 1999). For example, by constructing utilities “top down” through MAU analysis (using swing weights, for example) and comparing them to utilities built “bottom up” from choices (through conjoint analysis, a regression technique similar to bootstrapping) one can use discrepancies to reflect on why one’s “head” and “heart” disagree.

Few studies have attempted to examine the *validity* of weight-elicitation methods directly because of the lack of a natural accuracy criterion. A commonly used benchmark – holistic judgments of expected satisfaction – is a dubious choice given the weakness of holistic judgments in the linear models literature. The goal is to improve on holistic judgment. Stillwell et al. (1983) provided a clever strategy for overcoming this catch-22: They used output from a formal model that was familiar to the judges as a criterion for evaluating different weight-elicitation methods. More tools for verifying the effectiveness of MAU techniques are desirable.

Decision analysis (DA)

MAU is only one technique in the DA repertoire, which also includes methods for eliciting probabilities and for eliciting utilities under risk. Unfortunately, space does not permit reviewing each of these methods (see Pidgeon and Gregory, Chapter 30, this volume; Clemen, 1996, and Hammond, Keeney, & Raiffa, 1997, for excellent technical introductions; and Payne et al., 1999, for how DA techniques address decision flaws). At an abstract level, however, they share with MAU a set of common features. First, they all rest on an underlying logic of decomposition. D. Kleinmuntz (1990) has argued that all of these decomposition methods gain their effectiveness over holistic judgment by averaging out error in individual components (the same statistical principle that makes group judgments more effective than individual judgments). Second, DA methods rest on important *coherence* assumptions (about additivity, independence, etc.) that must be verified through a series of consistency checks. And, third, sensitivity analyses are important for determining whether conclusions are robust to different estimates of the components.

The quantitative nature of DA has traditionally made it a highly social process in which a technically astute advisor guides a decision maker through these procedures. This may create the impression of DA as complicated, expensive, and obscure. Recent approaches, however, have emphasized the qualitative aspects of *structuring* decisions, including identifying fundamental objectives and generating a broad set of new, creative alternatives (Hammond et al., 1997; Keeney, 2002). In addition, these authors have emphasized giving simpler versions of DA directly to individual decision makers (Hammond et al., 1997). Finally, there is an increasing potential to automate much of DA (discussed in the next section). All of these trends may help make DA more appealing and useful to ordinary decision makers. Unfortunately, perhaps the biggest shortcoming in debiasing research is the lack of empirical evidence on whether DA

training actually transfers to and improves everyday decisions – a shortcoming that begs to be addressed. Perhaps these new trends will provide more field and laboratory opportunities for studying decision analysis.

Decision Support Systems (DSS)

Computing technology has vastly enhanced the human ability to calculate and remember. Between the computer's capacity to execute complex algorithms in nanoseconds and to store libraries-worth of data, the computer can dramatically reduce the costs of effort in the *human* "effort-accuracy" tradeoff. In their useful summary of DSS, Edwards and Fasolo (2001, p. 581) speculate that computer-based "decision tools will be as important in the 21st Century as spreadsheets were in the 20th Century." But organizing technology around human tastes and limitations is still essential. As Edwards and Fasolo observe, "The idea of a procedure for making important decisions that does not depend heavily on human inputs seems unlikely as well as unattractive. Selection, training, and elicitation of responses from the person (or, more often, people) . . . become crucial" (p. 588).

DSS has the potential to improve individual decision making in a number of ways (although few systems currently exist that reflect all these possibilities):

- 1 DSS ensures the use of basic normative algorithms (MAU, Bayesian nets, Subjective Expected Utility Theory) that are otherwise hard for individuals to remember and to implement.
- 2 DSS can "bury" out of sight algorithms that would otherwise be intimidating to decision makers, making decision analytic tools more palatable.
- 3 DSS can run consistency checks (e.g., on probabilities or attribute weights) more easily and less obtrusively than a human advisor.
- 4 DSS can build and show the results of sensitivity analyses.

One of the most promising opportunities for improving decision accuracy through DSS is using information presentation to facilitate information acquisition and processing. Schkade and Kleinmuntz (1994) summarize the important dimensions of information display as:

- 1 the organization of displays (by alternative, by attribute, or as a matrix);
- 2 the form of displays (verbal or numeric); and
- 3 the sequence of information (sorted by preference or randomized).

In an extensive protocol study, they found that organization by alternative, attribute, or matrix led to corresponding differences in information *acquisition*, and that numeric displays yielded more compensatory *processing* than did verbal displays. And displays in which alternatives were sorted by decision maker's utility led decision makers to dwell on the most attractive options and to make faster decisions than did random displays.

Although the information displays did not influence choice quality in Schkade and Kleinmuntz (1994), it is reasonable to expect that factors that facilitate alternative-based acquisition and compensatory decision rules would yield superior decisions under many circumstances. Unfortunately, many current websites are not designed this way. Some websites offer huge data bases that allow a consumer to see a matrix of information and to sort alternatives by individual attributes. However, when there are scores of alternatives and dozens of attributes, consumers are essentially forced to resort to lexicographic decision rules. This is compounded by the fact that few sites provide the option of selecting on *several* attributes simultaneously, either by setting thresholds (a conjunctive decision rule) or by weighting and combining the attributes (in a compensatory way). Ideally, sites would attempt to capture a consumer's weights across attributes to facilitate compensatory tradeoffs in sorting alternatives (Edwards & Fasolo, 2001). The extent to which DSS and information displays can facilitate decision making is a growing area of study (Todd & Benbasat, 2000) that promises to become a central topic for debiasing research.

Adoption and Diffusion of Debiasing Techniques

One of the critical issues in debiasing research is identifying methods for facilitating the adoption of debiasing techniques. People resist being debiased for many reasons (Arkes, 2003; B. Kleinmuntz, 1990). They do not want to be told that they have been "doing it wrong" for all these years. They do not want to relinquish control over a decision process. And, perhaps most importantly, they fail to understand the benefits of many debiasing techniques relative to their own abilities, not just because they are overconfident, but because the techniques themselves are alien and complex, and the benefits are noisy, delayed, or small.

To understand the factors that promote adoption of a practice, it is useful to draw on a social psychological distinction between *compliance* to a behavioral norm and *internalization* of beliefs. Compliance is induced by rewards and coercion, and tends to produce superficial adoption – people mechanically go along with a practice in response to inducements, but abandon it when the inducements are removed. Internalization is fostered when a practice is endorsed by a trusted, expert source and when the acquisition process is active and voluntary, such as participating in the design of the practice (Kaplan, Reneau, & Whitecotton, 2001). Internalization is marked by an understanding of a practice and the intrinsic motivation to use it. Understanding, of course, is important for the successful adoption of cognitive strategies, especially for generalizing them across domains, but also for technological strategies, where ignorance about assumptions underlying DSS can be dangerous and overdependence can retard skills needed for *unaided* decision making (Glover, Prawitt, & Spilker, 1997).

With the exception of interesting research on factors that make decision aids more acceptable (e.g., Kaplan et al., 2001), there have been few studies on the adoption and diffusion of debiasing techniques, especially cognitive strategies. However, there are interesting case examples of the diffusion of similar practices. Every decade or so, a new

set of decision-making techniques seems to sweep through organizations, such as Total Quality Management (TQM) in the 1980s and Six Sigma in the late 1990s. Both movements included statistical tools, such as histograms and Pareto charts, in addition to softer tools, such as brainstorming and cause-effect analysis. Perhaps not surprisingly, the statistical tools are the first to be abandoned as practices diffuse through an organization (Zbaracki, 1998). The reasons that management fads die are many – senior management endorses something but does not practice it, engendering cynicism; people are skeptical of outside consultants brought in to train them; and trainees are separated from their coworkers, who put pressure on the trainee to continue with “business as usual” on the trainee’s return.

How might an organization overcome these obstacles? In the 1980s, Xerox took great pains to ensure the successful adoption of TQM by its 100,000 employees, including the statistical tools (Kearns & Nadler, 1992). They developed their own week-long training program and teaching materials. Existing organizational units were trained as a “family” and were expected to “learn it, use it, teach it, inspect it,” where “using it” included a mandatory “family” project to be done in the group’s home setting (thereby transforming declarative into procedural knowledge). The course was initially taught to the most senior managers, who then taught the managers at the next level below them; this process was repeated until the training had been “cascaded” down to the bottom of the organization. And teaching it required a deep understanding (if only to avoid embarrassment) as well as active, public endorsement of the methods. This behaviorally astute training program was perfectly designed to reduce cynicism, to encourage internalization, and to foster group support – not group resistance – for the new practices.

Heath et al. (1998) have argued that, in addition to studying formal training on formal techniques, it is also useful to understand the diffusion of informal debiasing techniques, which they called “cognitive repairs.” These feral debiasing strategies include a range of proverbs and procedures that are illustrated in Table 16.2. Examining cognitive repairs suggests several dimensions that affect the tendency for them to be adopted. The dimensions include: simple versus complex; domain-specific versus domain-general; social versus individual; and top-down versus bottom-up (which refers to where in an organization they originate). These dimensions come with inevitable tradeoffs. Many “cognitive repairs” in organizations are quite simple, involving a saying or an acronym, and domain-specific, tied to Wall Street or banks (as in Table 16.2). Together, simplicity and domain-specificity greatly enhance the memorability and applicability of the practices, making them more likely to be adopted in practice. But simple domain-specific repairs come at a cost. They tend not to be very precise (one would prefer a formula for discounting broker performance based on the mean and dispersion of the performance of other brokers) and they do not readily translate to new problems. But professional rules of thumb that point decision makers in the right direction are better than complex, general rules that are never understood, easily distorted, or quickly abandoned.

Heath et al. (1998) also argued that socially administered practices are often more effective than individual practices because individuals are overconfident in their decision-making abilities and fail to recognize when they need help. And, finally, Heath et al. argue that practices that emerge locally (or “bottom-up”) tend to have the advantage of

Table 16.2 Examples of “cognitive repairs”

<i>Cognitive repair</i>	<i>Benefits of the repair</i>
Wall Street brokers tell each other “Not to confuse brains with a bull market.”	This proverb helps deflate self-serving biases in decision-making ability.
Toyota and other companies encourage their employees to analyze problems by asking the question “Why?” five times.	The Five Whys helps decision makers to arrive at a deeper – rather than a merely accessible – answer for why problems have occurred.
At the Federal Reserve Bank of New York, examiners use an acronym known as CAMEL (capital adequacy, asset quality, management, earnings, and liability) to evaluate loans.	CAMEL ensures that examiners give attention to a full set of relevant attributes.
Motorola breaks up new product teams after the team has completed a project; subsequent new designs are pursued by newly formed teams.	By constantly rotating team members in a seemingly inefficient process, Motorola prevents the problem of shared views and shared errors discussed in the group decision making section.

simplicity, domain-specificity, and a sense of ownership. Combining the two, socially-administered practices that are homegrown, such as local norms of vigorous debate in scientific labs, tend to be more palatable than challenges imposed by superiors or outsiders. Heath et al. conclude that “the most successful repairs will be simple, domain-specific, socially administered, and evolved from bottom-up rather than developed from top-down. We find this conclusion intriguing because it describes repairs that differ sharply from those that are recommended in academic literatures on decision analysis, statistics, and economics” (pp. 30–1).

It is interesting to consider what implications these dimensions hold for the major classes of debiasing techniques reviewed in this chapter. Common socially-administered practices, such as incentives, accountability, and group decision making, guide decision makers to think more deeply than they would left to their own devices. Other practices can be internalized and used individually, such as “consider the opposite” or statistical and economic rules. However, these rules tend to have two strikes against them from an adoption perspective: They are often imposed “top-down” (in mandated statistics classes, for example) and they tend to be domain-general (impeding memorability and applicability). Their domain-generality is why actively applying rules to a broad range of examples is a critical feature of training on such rules. The hope is that they will be transformed from declarative to tacit knowledge – or, alternatively, they will “migrate” from System 2 to System 1 – as recognition of when and how to apply them becomes more automatic (see Kahneman, 2003; Phillips, Klein, & Sieck, Chapter 15, this volume).

The debiasing techniques that pose the greatest problem for adoption, however, are the technological strategies, such as statistical models or decision analysis, which are

complex, domain-general, and often imposed “top-down” by managers or consultants. Their logic may be difficult to understand for those less quantitatively trained, and their benefits are difficult to demonstrate in a vivid way, which is a classic obstacle to adoption. The possibility of transforming mysterious technological strategies into simpler, more intuitive, and more acceptable strategies is one of the great opportunities of debiasing research.

The Future of Debiasing

Research on debiasing tends to be overshadowed by research demonstrating biases: It is more newsworthy to show that something is broken than to show how to fix it. (It is tempting to propose that demonstrations of new biases must be accompanied by a debiasing technique, or at least a “moderator as repair” result.) However the sincere desire of many people in this field is to discover flaws not for their own sake, but with the intention of improving decision making. I have reviewed a number of effective debiasing techniques; more are needed. The development of new techniques will continue to be the central issue in debiasing research. But I hope that this chapter has also called attention to a central but neglected question in decision-making research: how do you encourage people to adopt better decision strategies?

I will close by speculating on two future directions for debiasing research. The first comes from the growing focus on how affect, motivation, and self-esteem influence decision making (see Larrick, 1993; Payne & Bettman, Chapter 6, this volume; Rottenstreich & Shu, Chapter 22, this volume). Identifying debiasing techniques for affect-based biases is a promising new area – What interventions help people cope effectively with emotion endogenous to a decision, such as anticipated regret? Or help them recognize and discount emotion that is extraneous to a decision, such as anger from some unrelated experience? The answers may bring decision research surprisingly close to clinical psychology, such as techniques used in cognitive-behavioral therapy. The second direction comes from a growing interest in the robustness of intuitive strategies (Gigerenzer, Chapter 4, this volume; McKenzie, Chapter 10, this volume; Phillips, Klein, & Sieck, Chapter 15, this volume). A future challenge for debiasing research will be assessing the benefits and costs of intervention: when is intuition sufficiently reliable that intervention is not worthwhile? Can decision makers be trained to recognize environments when they should trust their intuition and when they should modify or replace it (Hogarth, 2001; Payne & Bettman, Chapter 6, this volume)?

References

- Arkes, H. R. (1991) Costs and benefits of judgment errors: Implications for debiasing, *Psychological Bulletin*, 110, 486–98.
- Arkes, H. R. (2003) The nonuse of psychological research at two federal agencies, *Psychological Science*, 14, 1–6.

- Arkes, H. R., Dawes, R. M., & Christensen, C. (1986) Factors influencing the use of a decision rule in a probabilistic task, *Organizational Behavior and Human Decision Processes*, 37, 93–110.
- Brown, C. L. (1999) “Do the right thing”: Diverging effects of accountability in a managerial context, *Marketing Science*, 18, 230–46.
- Camerer, C. F. (1981) General conditions for the success of bootstrapping models, *Organizational Behavior and Human Performance*, 27, 411–22.
- Camerer, C. F. & Hogarth, R. M. (1999) The effects of financial incentives in experiments: A review and capital-labor-production framework, *Journal of Risk and Uncertainty*, 19, 7–42.
- Cheng, P., Holyoak, K. J., Nisbett, R. E., & Oliver, L. (1986) Pragmatic versus syntactic approaches to training deductive reasoning, *Cognitive Psychology*, 18, 293–328.
- Clemen, R. T. (1989) Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting*, 4, 559–84.
- Clemen, R. T. (1996) *Making Hard Decisions: An Introduction to Decision Analysis* (2nd edn.). Pacific Grove: Duxbury Press.
- Creyer, E. H., Payne, J. W., & Bettman, J. R. (1990) The impact of accuracy and effort feedback and goals on adaptive decision behavior, *Journal of Behavioral Decision Making*, 3, 1–16.
- Dawes, R. M. & Corrigan, B. (1974) Linear models in decision making, *Psychological Bulletin*, 81, 95–106.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989) Clinical versus actuarial judgment, *Science*, 243, 1668–74.
- Edwards, W. & Fasolo, B. (2001) Decision technology, *Annual Review of Psychology*, 52, 581–606.
- Fennema, M. G. & Kleinmuntz, D. N. (1995) Anticipations of effort and accuracy in multiattribute choice, *Organizational Behavior and Human Decision Processes*, 63(1), 21–32.
- Fischhoff, B. (1982) Debiasing. In D. Kahneman, P. Slovic, and A. Tversky (eds.), *Judgment Under Uncertainty: Heuristics and Biases* (pp. 422–44). Cambridge: Cambridge University Press.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986) The effects of statistical training on thinking about everyday problems, *Cognitive Psychology*, 18, 253–92.
- Fong, G. T. & Nisbett, R. E. (1991) Immediate and delayed transfer of training effects in statistical reasoning, *Journal of Experimental Psychology*, 120, 34–5.
- Gigerenzer, G. & Hoffrage, U. (1995) How to improve Bayesian reasoning without instruction: Frequency formats, *Psychological Review*, 102, 684–704.
- Glover, S. M., Prawitt, D. F., & Spilker, B. C. (1997) The influence of decision aids on user behavior: Implications for knowledge acquisition and inappropriate reliance, *Organizational Behavior and Human Decision Processes*, 72, 232–55.
- Griffin, D. & Buehler, R. (1999) Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology*, 38, 48–78.
- Hammond, K. R. & Summers, D. A. (1972) Cognitive control, *Psychological Review*, 79, 58–67.
- Hammond, J. S., Keeney, R. L., & Raiffa, H. (1997) *Smart Choices: A Practical Guide to Making Better Life Decisions*. New York: Broadway Books.
- Heath, C., Larrick, R. P., & Klayman, J. (1998) Cognitive repairs: How organizations compensate for the shortcomings of individual learners, *Research in Organizational Behavior*, 20, 1–37.
- Hogarth, R. M. (1978) A note on aggregating opinions, *Organizational Behavior and Human Performance*, 21, 40–6.
- Hogarth, R. M. (2001) *Educating Intuition*. Chicago: University of Chicago.
- Hogarth, R. M., Gibbs, B. J., McKenzie, C. R. M., & Marquis, M. A. (1991) Learning from feedback: Exactingness and incentives, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 734–52.

- Huber, O. & Seiser, G. (2001) Accounting and convincing: The effects of two types of justification on the decision process, *Journal of Behavioral Decision Making*, 14, 69–85.
- Jepson, D., Krantz, D. H., & Nisbett, R. E. (1983) Inductive reasoning: Competence or skill? *Behavioral and Brain Sciences*, 6, 494–501.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality, *American Psychologist*, 58, 697–720.
- Kahneman, D. & Frederick, S. (2002) Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (eds.), *Heuristics and biases* (pp. 49–81). New York: Cambridge University Press.
- Kaplan, S. E., Reneau, J. H., & Whitecotton, S. (2001) The effects of predictive ability information, locus of control, and decision maker involvement on decision aid reliance, *Journal of Behavioral Decision Making*, 14, 35–50.
- Kearns, D. T. & Nadler, D. A. (1992) *Prophets in the Dark: How Xerox Reinvented Itself and Drove Back the Japanese*. New York: Harper Business.
- Keeney, R. L. (2002) Common mistakes in making value trade-offs, *Operations Research*, 50, 935–45.
- Kleinmuntz, B. (1990) Why we still use our heads instead of formulas: Toward an integrative approach, *Psychological Bulletin*, 107, 296–310.
- Kleinmuntz, D. N. (1990) Decomposition and the control of error in decision-analytic models. In R. M. Hogarth (ed.), *Insights in Decision Making* (pp. 107–26). Chicago: University of Chicago Press.
- Larrick, R. P. (1993) Self-protective processes in decision making, *Psychological Bulletin*, 113, 440–50.
- Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990) Teaching the use of cost-benefit reasoning in everyday life, *Psychological Science*, 1, 362–70.
- Larrick, R. P., Nisbett, R. E., & Morgan, J. N. (1993) Who uses the normative rules of choice? Implications for the normative status of microeconomic theory, *Organizational Behavior and Human Decision Processes*, 56, 331–47.
- Lehman, D. R. & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate education on reasoning, *Developmental Psychology*, 26, 952–60.
- Leon, O. G. (1997) On the death of SMART and the birth of GRAPA, *Organizational Behavior and Human Decision Processes*, 71, 249–62.
- Lerner, J. S. & Tetlock, P. E. (1999) Accounting for the effects of accountability, *Psychological Bulletin*, 125, 225–75.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000) Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility, *Personality and Social Psychology Bulletin*, 26, 1142–50.
- Nisbett, R. E. (ed.) (1993) *Rules for Reasoning*. Hillsdale, NJ: Erlbaum.
- Nisbett, R. E., Krantz, D. H., Jepson, D., & Kunda, Z. (1983) The use of statistical heuristics in everyday reasoning, *Psychological Review*, 90, 339–63.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993) *The Adaptive Decision Maker*. Cambridge: Cambridge University Press.
- Payne, J. W., Bettman, J. R., & Schkade, D. A. (1999) Measuring constructed preferences: Towards a building code, *Journal of Risk and Uncertainty*, 19, 243–70.
- Raiffa, H. (1968) *Decision Analysis*. Reading, MA: Addison-Wesley.
- Schkade, D. A. & Kleinmuntz, D. N. (1994) Information displays and choice processes: Differential effects of organization, form, and sequence, *Organizational Behavior and Human Decision Processes*, 57, 319–37.

- Sedlmeier, P. (1999) *Improving Statistical Reasoning: Theoretical Models and Practical Implications*. Mahwah, NJ: Erlbaum.
- Siegel-Jacobs, K. & Yates, J. F. (1996) Effects of procedural and outcome accountability on judgment quality, *Organizational Behavior and Human Decision Processes*, 65, 1–17.
- Simonson, I. (1989) Choice based on reasons: The case of attraction and compromise effects, *Journal of Consumer Research*, 16, 158–74.
- Simonson, I. & Nye, P. (1992) The effect of accountability on susceptibility to decision errors, *Organizational Behavior and Human Decision Processes*, 51, 416–46.
- Soll, J. B. & Klayman, J. (2004) Overconfidence in interval estimates, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 299–314.
- Soll, J. B. & Larrick, R. P. (2004) Strategies for revising judgment: How, and how well, do people use others' opinions? Unpublished manuscript.
- Stanovich, K. E. (1999) *Who is Rational? Studies of Individual Differences in Reasoning*. Mahwah, NJ: Erlbaum.
- Stapel, D. A., Martin, L. L., & Schwarz, N. (1998) The smell of bias: What instigates correction processes in social judgments? *Personality and Social Psychology Bulletin*, 24, 797–806.
- Stillwell, W. G., Barron, F. H., & Edwards, W. (1983) Evaluating credit applications: A validation of multiattribute utility weight elicitation techniques, *Organizational Behavior and Human Performance*, 32, 87–108.
- Stone, D. N. & Zeibart, D. A. (1995) A model of financial incentive effects in decision making, *Organizational Behavior and Human Decision Processes*, 61, 250–61.
- Tetlock, P. E. & Boettger, R. (1989) Accountability: A social magnifier of the dilution effect, *Journal of Personality and Social Psychology*, 57, 388–98.
- Thaler, R. H. & Benartzi, S. (2001) Save more tomorrow: Using behavioral economics to increase employee saving. Working paper, University of Chicago.
- Todd, P. & Benbasat, I. (2000) Inducing compensatory information processing through decision aids that facilitate effort reduction: An experimental assessment, *Journal of Behavioral Decision Making*, 13, 91–106.
- Tversky, A. & Kahneman, D. (1983) Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment, *Psychological Review*, 90, 293–315.
- Wegener, D. T. & Petty, R. E. (1995) Flexible correction processes in social judgment: The role of naïve theories in corrections for perceived bias, *Journal of Personality and Social Psychology*, 68, 36–51.
- Wilson, T. D. & Brekke, N. (1994) Mental contamination and mental correction: Unwanted influences on judgments and evaluations, *Psychological Bulletin*, 116, 117–42.
- Wilson, T. D. & Schooler, J. W. (1991) Thinking too much: Introspection can reduce the quality of preferences and decisions, *Journal of Personality and Social Psychology*, 60, 181–92.
- Zbaracki, M. J. (1998) The rhetoric and reality of total quality management, *Administrative Science Quarterly*, 43, 602–36.