1   Hard Decisions, Bad Decisions: On Decision Quality and Decision Aiding

J. Frank Yates, Elizabeth S. Veinott, and Andrea L. Patalano

ABSTRACT

Behavior-focused decision aids have had little documented success. A proposed contributor is this: To most deciders, *decision quality* entails myriad diverse facets, with an emphasis on material welfare. Yet, the typical decision aid (and its theoretical underpinning) is predicated on a narrow conception of decision quality that has other emphases. Deciders therefore often ignore such aids because they appear irrelevant to significant decider concerns. And when deciders do try the aids, the results disappoint them because the aids leave untouched quality dimensions that matter to them. Two empirical studies and a critical review of the most popular aiding approaches (from decision analysis to expert systems) support this thesis. The chapter offers for consideration a new, comprehensive decision quality conception intended to facilitate both fundamental and practical scholarship. The analysis also argues for *decision attribution theories* that would explain how deciders think they decide and why they believe that their decisions sometimes fail.

CONTENTS

How can we help people make better decisions? This question has inspired the labors of virtually all decision behavior scholars, including those who focus their attention on *fundamentals* (e.g., whether, how, and why choices among gambles violate the axioms of expected utility theory). Unfortunately, in many practical situations, there is little hard evidence that the techniques and devices, that is, decision aids, growing from these efforts have, in fact, yielded substantial, demonstrable improvements in how people decide (e.g., O'Connor et al., 1999). And consultants who have made valiant efforts to promote the application of behavior-focused aids in such settings privately acknowledge that prospective users of the aids have been increasingly indifferent – if not hostile – to those tools. Critics of traditional decision scholarship identifying themselves with the *naturalistic decision-making* movement (cf. Klein, Orasanu, Calderwood, & Zsambok, 1993; Zsambok & Klein, 1997) have conveyed the same message, but more openly and bluntly. Many of these critics have actually tried the tools, found them wanting, and been left disillusioned.

Why have behavior-focused decision-aiding efforts so often met with minimal success? There are undoubtedly many reasons for this state of affairs; failure frequently has many parents. Here, however, we concentrate on just one particular potential contributor, which turns on the concept of *decision quality*. The plan of this chapter is as follows: In the first section, we present and elaborate our focal proposition, that decision quality is not a unitary construct in people's heads and that decision aids often fail because they do not address the dimensions of decision quality that people wish to improve. The second section describes two

by *quality?* The thesis we entertain is this: In the mind of the typical person – be that person a decider or a decision scholar – decision quality is not a unitary construct. Instead, it consists of several distinct, imperfectly correlated facets. Moreover, the facets comprising one person's notion of decision quality are unlikely to correspond perfectly to those of another, and indeed may well surprise that other person. (*"That's* what you mean by a good decision?") Further, the facets that are – or are seen by deciders to be – most pertinent in some decision problems can be markedly different from those most relevant in other problems. And, finally and significantly, the quality conceptions implicitly assumed and addressed in any specific decision-aiding effort (or body of decision scholarship) are almost necessarily incomplete relative to the full range of facets that legitimately comprise any single person's notion of decision quality. These propositions, if true, would constitute a partial account for the apparently limited success of such aiding efforts. Suppose that the decider is concerned about, say, quality facets F1, F2, F3, F4, and F5. According to the present thesis, there is a good chance that any decision aid offered to the decider will appear to address only, say, quality facets F2 and F6, where the decider sees F6 as entirely irrelevant. Little wonder, then, that the decider would find that aid unappealing and would not even try it. Any actual potential of the aid for enhancing decision quality would be moot.

Empirical Studies: Decision Quality Facets for Personal Decisions

Our basic thesis was motivated by extensive study of the literatures on decision making in numerous disciplines. It was also grounded in qualitative observations and informal interviews of seemingly countless deciders discussing their real-life decision problems. Here we describe two empirical studies we performed in order to test the impressions originating in those less structured inquiries. In the best of circumstances, such studies would examine decisions made by people in a wide variety of contexts. But practical constraints precluded that. Thus, we focused on personal decisions made by convenience samples of college students. The specific considerations identified in our studies should in no sense be regarded as representative. Yet, there appears to be no reason to expect that the broad categories of decision quality facets implicated by those considerations do not apply to decision problems and deciders generally.

empirical studies that provide evidence bearing on the plausibility of that proposition. The third section sketches and characterizes a representative sample of leading behavior-focused decision-aiding strategies. We show how those characterizations, coupled with the results of the empirical studies, buttress the claim that problematic aspects of decision quality conceptions really do play a significant role in determining how well decision-aiding attempts fare. In the fourth and final section of the chapter, we discuss the implications of our analyses for decision aiding and for decision scholarship more generally.

The Decision Quality Thesis

The following definition of a *decision* is a synthesis of how the term is actually understood and used across the myriad disciplines that study decision making, not just psychology (cf. Yates & Estin, 1998; Yates & Patalano, 1999):

> **A decision is a commitment to a course of action that is**
> **intended to produce a satisfying state of affairs.**

Thus, quality is part and parcel of the very idea of a decision. Consider, for instance, a decision that yields a more satisfying state of affairs for the implied beneficiary (who may or may not be the person making the decision) than does some other decision. This is equivalent to saying that that first decision has higher quality than its competing alternative. Or consider the "better decisions" sought by decision scholars more generally. These are, implicitly, ones that have quality superior to that possessed by "worse decisions."

The decision idea further suggests that when a decision maker – a *decider* – is confronted with a decision problem, a metadecision process focusing on decision quality ensues. In the voice of the decider, the following kind of soliloquy can be expected on some occasions:

> Suppose I were to make this decision the way that I'm naturally inclined to make it. What measure of quality could I expect? . . . That bad, huh? It sure would be good if I had help assuring better quality.

Hence, on such an occasion, the decider would be receptive to a decision aid that appears to enhance the chances of making a decision with adequately high quality. But what, exactly, does the decider mean

Study 1: Decision "Hardness" and "Easiness"

Two major classes of decision quality facets can be distinguished. The first class consists of aspects associated with the *products* of a decision per se. That is, they concern what the decision yields for the decision beneficiary. In contrast, the second class of decision quality aspects concerns the *process* by which a decision is made, more specifically, *decision process difficulty* as experienced by the decider. In other words, they pertain to what makes a given decision problem either "hard" or "easy" for the decider to solve. It is legitimate to regard difficulty as an element of decision quality for the same reason that people consider the difficulty of negotiating for a new car to be part of the overall adequacy of the transaction. Study 2 was designed to be enlightening about product aspects of decision quality. In contrast, Study 1 focused on decision process difficulty, a kind of decision-making *cost*, in the broad sense. One justification for using a decision aid would be the expectation that it would alleviate the difficulties associated with arriving at a satisfactory decision (cf. Clemen, 1991). The objective here was to discover the specific *kinds* of difficulty that people experience generally and hence plausibly would seek to avoid when they are in situations where they have to decide. That people would indeed welcome such relief is implicit in the following kind of remark often reported in news accounts of both heroic and tragic events: "Sadly, we had to make a very tough decision."

Method

Participants. Ninety-nine introductory psychology students at the University of Michigan participated in this study in exchange for course credit. Sixty-three percent were female.

Procedure. A questionnaire in booklet form was administered to participants in groups of 10–15 people. Participants were instructed to "imagine hard and easy decisions that you have made." Each participant was then asked to describe two or three hard decisions and two or three easy ones that "you have made in the past year." Since hard decisions had priority for us, out of concern for time constraints, all participants wrote about their hard decisions first. For both hard and easy decisions, participants were asked three key questions (among others). They were asked to first describe the circumstances that gave rise to each decision problem, then to explain why the decision was hard (easy) for them, and after that to indicate how they solved the given problem.

Further, for each decision, each participant was asked to report (a) how many options had been available; (b) how long it took to make the decision; (c) the degree to which the participant felt that he or she took either too little or too much time making the decision; (d) how satisfied the participant was with the decision's eventual outcome; and (e) the extent to which the participant felt that his or her decision was the best possible given the circumstances.

Results and Discussion

Basic hard/easy comparisons. In total, participants described 212 hard decisions and 200 easy ones. The topics of these decisions ran the gamut but tended to fall into three major categories: academic issues (e.g., which college to attend, what major to choose, which classes to take), relationships and social life (e.g., whom to date, which fraternity or sorority to join), and financial matters, such as consumer purchases (e.g., which car to buy). For both hard and easy decisions, participants typically reported choosing among three or four options. Several pertinent hard versus easy comparisons included the following:

- *Decision time:* Hard decisions took a median of 3 weeks to make, whereas easy decisions typically took only 2 days. Although participants spent more time making hard decisions than easy ones, there was no difference in their ratings of the appropriateness of the amount of time taken, $t(86) = -1.90$, *ns*. That is, participants appeared to feel that hard decision problems simply demanded more time to solve properly and that they gave those problems their due. Of course, it is possible that hard and easy problems tend to have different deadlines attached to them, too.
- *Satisfaction:* Participants' satisfaction with the outcomes of their decisions was rated on a 9-point scale, where 1 = "Not At All Satisfied" and 9 = "Extremely Satisfied." The mean satisfaction ratings for hard and easy decision outcomes were 7.6 and 8.1, respectively, $t(89) = 2.71$, $p < .01$, for a within-participants comparison on the first hard and the first easy decision cited by each participant with complete satisfaction data. That is, participants were significantly more pleased with the results of their easy decisions than their hard decisions. Note, however, that the magnitude of the difference in satisfaction ratings was small, only half a scale point. And it is

especially noteworthy that the mean rating for the hard decision outcomes was less than one and a half scale points from the maximum possible. In other words, on average, no matter what, participants were rather pleased with what their decisions yielded.

- *Relative decision adequacy:* One defensible (and common) definition of a *good decision* is that it is the selection of the best alternative available at the time the decision is made. This conception of good decision making is useful because it acknowledges that, in some circumstances, all potential outcomes are unpleasant in an absolute sense (e.g., in medical situations where patients are already in irreversibly poor health). A measure of decision adequacy that focuses solely on absolute ratings of outcome satisfaction would neglect such possibilities. That is why participants here were also asked to rate the adequacy of their decisions relative to, in hindsight, the best available on a 9-point scale, where $1 =$ "Not the Best" to $9 =$ "Definitely the Best." The mean ratings for hard and easy decisions were, respectively, 7.6 and 8.0, $t(91) = 2.06$, $p < .05$. Once again, although participants were less pleased with their hard decisions than with their easy ones, by no means did they regret either.

"Why hard/easy?" coding procedures. Our primary aim was to understand how people come to regard some decisions as hard but others as easy. Implicit is the assumption that subjective hardness is simply the lay characterization of a decision being difficult to make and that hardness attributions are specifications of the different kinds of difficulty that weigh upon deciders. Recall that our data-collection procedure was to have participants bring to mind decisions that they themselves had made and that they themselves classified as either hard or easy. Participants then gave their own accounts of why they made those classifications. The reported explanations constituted our basic data. Consistent with our initial proposition, participants' explanations for what made decisions either hard or easy were remarkably diverse. Nevertheless, it was obvious that there was structure within that diversity.

Our strategy for discerning the structure underneath the data and then encoding them for further analysis was the following: First, each of the investigators independently read the protocol for every decision

described by every participant. Each investigator then developed his or her own scheme for classifying the participants' reasons for considering a decision to be either hard or easy. After that, the investigators reached a consensus about a common coding scheme for subjective hardness and another for subjective easiness. The hardness scheme contained 29 categories and the easiness scheme included 24.

With these coding schemes in hand, we then developed training materials and taught two naive coders to apply the respective schemes to the original data. The coding procedure required that the coder consider explicitly whether each category in the coding scheme was or was not present in a given protocol. Thus, it was entirely possible that a given decision might have been regarded as hard (or easy) for several reasons, not just one. (Ultimately, it turned out that the number of hard codes assigned to the typical hard decision was 1.9, and the average easy decision was given 1.6 easy codes.) The coders were instructed to be conservative, to encode only direct attributions for why a decision was hard or easy, not to infer what attribution "made sense" given the participant's remarks or the situation described. For example, buying a car might seem like a big expense for anyone. But unless the participant discussing a car-buying decision explicitly mentioned that expense was a reason the decision was hard, the coder was not to record the "expense" hardness category. After training, each coder independently encoded every protocol. At regular intervals as the coding activity proceeded, the coders met and compared their codes. When discrepancies arose, the coders discussed their disagreements and reached a final consensus about whether a given category was or was not represented in a particular protocol.

Emergent hardness categories. The hardness and easiness categories are most usefully interpreted when they are structured into a smaller number of *supercategories*. One particularly enlightening structure has seven supercategories for hardness. A parallel structure was recognized for easiness as well. Here we describe the supercategories encompassing the 29 hardness categories that emerged from the data in the kind of language a decider might use in characterizing a given form of hardness:

- *Hardness 1 – Outcomes: Serious: "This decision is hard because a serious loss of some kind can (or is likely to) result from it."* The specific loss categories commonly cited by participants included ones

with long-term, possibly irreversible, effects, ones that entailed hurting another person, ones that required violating personal (e.g., moral) principles, ones that involved large, significant outcomes, and ones that held great risks.

- *Hardness 2 – Options: "This decision is hard because of burdens imposed by the number and/or character of the available options."* Thus, there might be too many or too few options among which to choose, or those options might require comparisons on too many factors.

- *Hardness 3 – Process: Onerous: "This decision is hard because the process of making it is onerous."* Among the specific kinds of onerousness cited by the participants were the amount of effort required, the presence of emotional circumstances, time pressure, uncertainty, and the decider's feeling that he or she lacks essential expertise.

- *Hardness 4 – Possibilities: "This decision is hard because it is difficult to imagine or predict what its possible outcomes might be."* One particular form this variety of hardness takes arises when the decider has had little or no experience with the kinds of alternatives under consideration.

- *Hardness 5 – Clarity: "This decision is hard because it is especially unclear which alternative is superior to its competitors with respect to the considerations on which they are being compared."* An important variant of clarity hardness is simply a lack of dominance, that is, there is no option that is at least as good as every other option with respect to every consideration and better with respect to at least one of those considerations. Cases of "tying" implicate another variant, where, in the aggregate, two or more options seem tied for first place in their overall appeal; for example, the decider would be inclined to say, "But I love them both!"

- *Hardness 6 – Value: "This decision is hard because I am unsure how I would feel about specific outcomes that might result from it."* For instance, the decider might have never experienced some particular outcome and is uncertain whether the experience would be pleasant or unpleasant and to what degree.

- *Hardness 7 – Advisors: "This decision is hard because of conflicting recommendations or advice."* In one kind of situation where this type of hardness occurs, the decider is faced with advisors who contradict one another. In another, an advisor who is important

to the decider pushes views or actions that contradict the decider's own opinions.

Emergent easiness categories. The supercategories for easiness were essentially mirror images of the hardness supercategories. They can be characterized as follows:

- *Easiness 1 – Outcomes: Mild: "This decision is easy because its potential outcomes are insignificant, nothing to worry about."* Specific instances of problems involving this kind of easiness entailed considerations such as short-term effects, reversibility, minimal risk, and the existence of a win-win situation with options that precluded any sort of loss.

- *Easiness 2 – Options: "This decision is easy because the available options require minimal reflection."* One variant of this type of easiness occurs when a decision must be made in order to achieve a specific objective (e.g., satisfying a requirement that a student must complete an advanced laboratory course to earn a particular degree) and there is indeed an option available that clearly attains that goal. In a decision entailing another variant, the constraints are so exacting that otherwise appealing alternatives are eliminated out of hand, leaving only a small, manageable number to ponder.

- *Easiness 3 – Process: Benign: "This decision is easy because the process of making it is not unpleasant."* One common form of process easiness was identified when participants said that they "just knew" which option to pick. Others were indicated when participants said that their decisions took minimal effort, that they had had experience making similar decisions before, or when one option was favored with respect to some overriding consideration.

- *Easiness 4 – Possibilities: "This decision is easy because projecting its possible outcomes is straightforward."* Decisions involving alternatives that participants had experienced previously were often easy in the possibilities sense; the decider felt sure about what *could* happen, even if he or she did not know what *would* happen if any particular option were pursued.

- *Easiness 5 – Clarity: "This decision is easy because it is readily apparent which alternative is better than its competitors, taking into account all the pertinent considerations."* A common form of this kind of easiness arose when, for whatever reason, the available

option pool contained one dominating alternative, for example, it had all positive features and no negatives. In another form, one option was clearly closer to the decider's ideal than were its rivals.

- *Easiness 6 – Value:* "*This decision is easy because I know for sure how I would feel when experiencing each of its potential outcomes.*" Such confidence generally arose among participants as a result of their prior experience with those outcomes, which made clear to them whether they liked or disliked those outcomes and to what extent.
- *Easiness 7 – Advisors:* "*This decision is easy because of the recommendations or encouragement of others.*" One form such easiness can take entails essentially turning the decision problem over to others, letting them figure out what the decider ought to do.

Category incidence rates. Table 1.1 displays the percentages of decisions the participants classified as hard and easy for reasons belonging to the various hardness and easiness supercategories just described. Several conclusions are indicated by the findings summarized in the table. The first is that there are indeed many distinct kinds of decision difficulty that weigh upon deciders; there was considerable breadth and variety in the ways our participants saw their decisions as being hard or easy. A second conclusion implicit in the table is that hardness and easiness citations are not complementary. That is, even if a particular category was cited often as a reason hard decisions might be hard, this provided no assurance that the parallel category would be cited just as frequently as a reason for easy decisions being easy. And then there are the specific categories that were most commonly mentioned. Particularly notable is "Outcomes: Serious Loss Potential" for hardness. The data indicate that, by far, when people say that a decision is hard, they mean that they are concerned about things like the stakes involved – what the decision beneficiary could lose through the decision.

At least tentatively, the data have several implications for those seeking to provide decision-making assistance that would be well received by prospective clients. Most obviously, the description (and, ideally, the reality) of that assistance should convince the client that that assistance would reduce significantly the client's chances of experiencing a serious loss of any kind. (The implied hypersensitivity to losses is reminiscent of

Table 1.1. *Percentages of Decisions Classified by Participants at Least Once as Hard or Easy for Reasons Corresponding to Various Hardness and Easiness Supercategories*

Supercategory	Percentage
Outcomes	
Hard (Serious loss potential)	69.8%
Easy (Insignificant)	17.5%
Options	
Hard (Too many/few, Character)	10.8%
Easy (Minimal reflection required)	21.0%
Process	
Hard (Onerous)	22.1%
Easy (Benign)	31.0%
Possibilities	
Hard (Obscure)	8.0%
Easy (Apparent)	15.0%
Clarity	
Hard (Ambiguous superiority)	23.1%
Easy (Obvious superiority)	41.5%
Value	
Hard (Uncertain)	7.5%
Easy (Clear-cut)	21.5%
Advisors	
Hard (Disagree)	5.1%
Easy (Recommend, Encourage)	4.5%

Note: Hard and easy supercategory percentages sum to more than 100% because participants typically cited more than one reason that a given decision was hard or easy.

numerous assertions and results in the decision behavior literature, such as the steepness of value functions for losses postulated in Kahneman and Tversky's 1979 prospect theory.) The other supercategories with sizable incidence rates (e.g., 20% or more) contain advice that might well be wise, too. Thus, for instance, a decision aid developer and consultant would probably find it fruitful to build and promote aids that are transparently effortless and perhaps even fun to use (addressing deciders' process concerns). Those aids should also make one of the decider's options stand out as distinctly better than its competitors (addressing deciders' implicit demands for clarity). And, if at all possible, an aid should help the decider to clearly and, presumably, accurately anticipate how the decider would actually feel about the potential outcomes

of the actions under consideration (speaking to deciders' sensitivity to value clarity).

Study 2: Decision "Badness" and "Goodness"

Recall that Study 1 was intended to shed light primarily on people's conceptions of decision quality most closely associated with the difficulty of the processes by which they arrive at their decisions. In contrast, Study 2 was designed to illuminate quality notions identified more strongly with the products of decisions. Our guiding assumption was that these notions are especially likely to be manifest when people reflect upon real decisions they have made and that they themselves regard as having been either good or bad. Thus the following approach, which paralleled the one used in Study 1.

Method

Participants. One hundred and ten introductory psychology students at the University of Michigan participated in this study in exchange for course credit. Fifty-four percent were female.

Procedure. Each participant completed a questionnaire that was administered by computer. The questionnaire first told the participant that our purpose was to understand what participants thought made a decision either good or bad. The participant was told that we would focus on decisions the participant had actually made, but only decisions about which the participant had had to "think hard." The computer then required the participant to bring to mind four such decisions made within the previous year, two good decisions and two bad ones. For each of those decisions, the participant first wrote a brief descriptive title. The participant then rated the goodness or badness of the decision "relative to all the important decisions you have ever made" on an 11-point scale ranging from −5 ("Extremely Bad") to 0 ("Neither Good Nor Bad"), to +5 ("Extremely Good"). Next, the participant rated the importance of the decision, again "relative to all the important decisions you have ever made," on an 11-point scale ranging from 0 ("Not Important At All") to 10 ("Extremely Important"). The product of these quality and importance ratings yielded an *impact score* for the decision. The computer then selected two of the participant's four decisions for further consideration, the "really bad" decision with the more negative impact score and the "really good" decision with the more positive impact score.

The computer next requested that the participant answer a series of questions about each of the focal good and bad decisions. (The good–bad and bad–good orders were counterbalanced across participants.) First, the computer asked the respondent to "explain why you classified the present decision as a bad (good) one." It then asked for specific details about how the decision was actually made, including when and how the decider came to realize that there was a decision to make and how long it took to make that decision.

Basic bad/good comparisons. As implied by the preceding procedure description, 110 bad and 110 good decisions were examined, one of each for every participant. The mean quality ratings of the focal good and bad decisions were, respectively, +3.6 and −2.4 (on the scale from +5 to −5), $t(109) = 28.99$, $p < .001$, thus providing something of a check on the manipulation. Observe, however, considering the distance of the ratings from the 0 neutral point, that the focal good decisions seemed to be better than the focal bad decisions were bad. The mean importance ratings of the focal good and bad decisions were 7.7 and 5.6, respectively, $t(109) = 7.22$, $p < .001$. That is, the participants' bad decisions did not seem as important to them as their good ones. Taken together, the quality and importance ratings suggest the same story indicated in Study 1: People are, on the whole, not greatly displeased with the real-life decisions they make. Perhaps substantively significantly (a point to which we return later), on average it took participants far less time to bring to mind their bad decisions (53 seconds) than their good ones (70 seconds), $t(109) = 4.78$, $p < .001$. Nevertheless, they required about the same amount of time to write about them, approximately 10.5 minutes, $t < 1$.

"Why bad/good?" coding procedures. We used basically the same approach as in Study 1 for developing and applying schemes for encoding participants' explanations for why their decisions were either bad or good. So there is no need for us to review the procedural details. We should note, however, that those procedures yielded 20 categories of decision badness and 24 of decision goodness. Further, on average, 2.3 distinct coding categories were assigned to each bad decision and 3.7 to each good decision. In and of itself, as was the case with perceptions of hardness and easiness, this constitutes evidence of the multifaceted character of people's subjective notions of decision quality. Decisions are not seen as simply good or bad to some degree; they excel or fall short with respect to a host of qualitatively distinct dimensions that capture people's attention.

Emergent badness categories. The supercategory structures for badness and goodness contained five parallel supercategories each. We first describe the supercategories for badness in the voice of a decider explaining why he or she regarded a particular decision as bad.

- *Badness 1 – Experienced Outcomes: Adverse: "This decision was bad because it resulted in bad outcomes."* Recognition of violating personal (e.g., moral) principles was one major form this kind of badness assumed. Another entailed any sort of bad outcome that had never even been contemplated at the time the decision was made, that is, *blindsiding.* A related but different variety was implied when there were bad outcomes that had been recognized as possible but were not actually *expected* to occur.

- *Badness 2 – Missed Outcomes: Favorable: "This decision was bad because it resulted in me missing out on good outcomes."* In one form of this kind of badness, the decision caused the decider to miss out on some good experience that would have occurred otherwise. In another, the decision led to the relinquishment of a good thing the decider already possessed, such as a satisfying relationship.

- *Badness 3 – Options: "This decision was bad because of its implications for my options, presently or in the future."* In one common form of this kind of badness, the decision limited the decider's future options, such as closing off certain career paths. When participants cited a second, "fuzzier" variety, they reported that they knew that there were better options than the ones they chose but, for no good reason, they rejected those options.

- *Badness 4 – Process: "This decision was bad because the process used to make it was bad."* This process category did not seem to be differentiated in any principled way. Participants who cited factors that fell into this category felt that, for miscellaneous reasons, such as minimal care or time devoted to the task, and independent of its actual outcomes, the process employed in making a decision was flawed.

- *Badness 5 – Affect: "This decision was bad because I felt bad while (or after) making it."* A sense that the decision exposed the decider to risk was reported for one variation of this kind of badness (e.g., the risk of contracting a sexually transmitted disease). The feeling of regretting an action defined another.

Emergent goodness categories. These were the supercategories for goodness, again as a decider might characterize them.

- *Goodness 1 – Experienced Outcomes: Favorable: "This decision was good because it yielded good outcomes."* For the most frequently cited version of this first form of goodness, participants indicated that the given decision had already produced good outcomes. In another, however, such outcomes were merely anticipated. A third variation acknowledged that, although not all decision outcomes were favorable, on balance, those outcomes *tended* to be good ones. In still another form, which we labeled *relative elation*, the decider believed that the decision's outcomes were better than those that would have resulted from another option the decider might have selected but did not.

- *Goodness 2 – Missed Outcomes: Adverse: "This decision was good because it prevented me from experiencing bad outcomes."* The first variety of this kind of goodness occurred when the decision precluded the decider from experiencing bad outcomes that otherwise would have taken place. The alternative form arose when the decision rescued the decider from a bad situation that already existed, such as a dysfunctional relationship.

- *Goodness 3 – Options: "This decision was good because of how it improved my options."* When this type of goodness was cited, making the pertinent decision opened up new (and presumably attractive) alternatives that either did not exist before or were at least unrecognized.

- *Goodness 4 – Process: "This decision was good because the process used to make it was good."* When participants reported this category of goodness, they mentioned that some aspect of how the decision was made – irrespective of its outcomes – contributed to its appraisal as a good decision.

- *Goodness 5 – Affect: "This decision was good because I felt good while (or after) making it."* One variation of affect goodness entailed a general sense of pleasure with the given decision's outcomes. The other most common form of this kind of goodness occurred when the decision made the decider feel good about him- or herself, for example, experience a sense of pride.

Category incidence rates. Table 1.2 presents the percentages of decisions the participants regarded as bad and good for reasons belonging to the badness and goodness supercategories just sketched. The table

Table 1.2. *Percentages of Decisions Classified by Participants at Least Once as Bad or Good for Reasons Corresponding to Various Badness and Goodness Supercategories*

Supercategory	Percentage
Experienced outcomes	
Bad (Adverse)	89.0%
Good (Favorable)	95.4%
Missed outcomes	
Bad (Favorable)	25.7%
Good (Adverse)	30.3%
Options	
Bad (Limiting)	44.0%
Good (Improved)	14.7%
Process	
Bad (Flawed – outcome independent)	20.2%
Good (Sound – outcome indpendent))	6.4%
Affect	
Bad (Risk exposure, Regret)	23.9%
Good (Pleasure, Self-esteem)	40.4%

Note: Bad and good supercategory percentages sum to more than 100% because participants typically cited more than one reason that a given decision was hard or easy.

clearly buttresses the conclusion that deciders' personal decision quality concepts are multifaceted, entailing constructs that are typically quite distinct from one another. Implicit in the data is also the conclusion that the acknowledged facets tend to differ from person to person and from one decision to the next. An especially striking feature of the results is that subjective notions of decision quality are overwhelmingly dominated by outcomes: Good decisions produce good outcomes and bad decisions yield bad ones, directly or in the opportunity cost sense. We are also struck by the extent to which people's appraisals of decisions are affected by abstractions, for example, the kind of counterfactual reasoning implicit in the "Missed Outcomes" supercategories.

As with Study 1, it is useful to consider the practical implications of the incidence rates in Table 1.2 for those who aspire to help people decide better. Again, but in even more dramatic fashion, the data (concerning the outcome supercategories) indicate that a decision aid developer or consultant simply *must* convince clients that his or her offerings have a good chance of improving clients' situations materially and even morally. The options responses suggest that it would be

wise to speak to deciders' concerns about the long-term implications of their decisions as well. And the affect citations imply that decision aids and consultations ideally should seek to have deciders feel an undifferentiated "warm glow" about the decision-making experience and its aftermath as well. Further, at least some attention should be devoted to conveying the idea that an aid or consultation is sound in ways that do not translate immediately and directly into favorable outcomes for the decision beneficiary.

A Review: Quality Conceptions Implicit in Decision-Aiding Practice

There are many approaches to behavior-focused decision aiding as well as specific aids derived from those approaches. Nevertheless, they fall into a relatively small number of major classes. Here we briefly describe six of the most popular ones: decision analysis, debiasing techniques, social judgment theory (and its relatives), general decision support systems, group decision support systems, and expert systems. In each case, we attempt to discern the conceptions of decision quality that developers and practitioners have (implicitly or explicitly) sought to address and enhance with their techniques and devices. We then try to reconcile the foci of those aids with what our empirical studies suggest that deciders think that they *need* to address in order to make adequate decisions.

Decision Analysis

At one level, at least, decision analysis is an exceptionally broad and comprehensive approach to decision aiding. Applications have been reported in domains as diverse as the siting of public utilities (e.g., Keeney & Nair, 1977), capital investment in mining (e.g., Hax & Wiig, 1977), policies for treating intracranial aneurysms (Aoki et al., 1998), and mediation in legal trials (e.g., Aaron, 1995). Von Winterfeldt and Edwards (1986) wrote a book on decision analysis, which they introduced as "a set of formal models, semiformal techniques for interacting with people, and bits of lore and craft" (p. 2) used in the service of helping people to be rational in making inferences and decisions. They described the goal of decision analysis as structuring and simplifying the task of making hard decisions as well and as easily as the nature of those decisions permits. Decision analytic methods include such things as (a) ways to organize or structure a decision problem (e.g., with decision trees or

influence diagrams), (b) techniques for assessing uncertainty (including the use of debiasing techniques, discussed subsequently), (c) procedures for measuring value and utility, (d) operations for combining information to arrive at a choice (e.g., expected utility theory and multiattribute utility theory), and (e) sensitivity analyses (discerning how much conclusions depend on the precise values used). Decision analysts are not greatly concerned with inherent bias in human strategies per se. The assumption is that, as long as time and effort are taken to elicit biases in multiple ways, various techniques (including debiasing procedures) can be applied to ensure that reasonable judgments are acquired to guide decisions properly. Furthermore, decision analysis uses tools such as sensitivity analysis to ensure that small judgment differences do not significantly affect the recommendations derived in an analysis.

As suggested by von Winterfeldt and Edwards's (1986) description of the field, typically, as far as decision analysts are concerned, decision quality is embodied in the rationality or logicalness of the process by which people decide. And the particular variety of rationality of interest is commonly referred to as *internal consistency* or *coherence* (cf. Yates, 1990, chapters 5 and 9). When a person is rational in the coherence sense, that person avoids endorsing principles that contradict one another. That person also seeks to avoid holding beliefs and making choices that conflict with principles he or she accepts as appropriate, for example, the axioms of probability theory, expected utility theory, or a particular variety of multiattribute utility theory. Implicitly – and sometimes explicitly – analysts assume that adherence to rationality is threatened by two recurrent features of decision situations: uncertainty and multiple conflicting objectives. Hence the character of the techniques described previously. Our empirical results suggest that real deciders rarely acknowledge a concern with the rationality of their decision processes and certainly not in the coherence sense. Instead, they are preoccupied with results – good outcomes of various kinds. Analysts sometimes imply that there is in fact a strong link between coherence and results (e.g., Morris, 1977, p. 13). And some appear to believe that the typical decider shares that faith, as suggested by the remarks of Matheson and Howard (1968, p. 12): "Most persons follow logical decision procedures because they believe that these procedures, speaking loosely, produce the best chance of obtaining good outcomes."

Our data do not prove the point, but they are consistent with the possibility that faith in a strong coherence–outcome link is *not* widely shared. If real, this lack of faith would constitute a potentially significant

contributor to the relative unpopularity of decision analysis. As we will discuss, this is not to say that people think that decision rationality and outcomes are independent. Instead, people might simply believe that rationality, in the decision analytic, coherence sense, is far from sufficient for producing decisions that yield good outcomes. And the obvious demands and burdens imposed by those methods (e.g., time and tedium) lessen their appeal even more. Of course, analysts might well respond that decision analysis is simply inappropriate for the small-scale personal problems of students like our research participants; the stakes do not warrant the required investment. (On the other hand, at least some decision analysis advocates would probably disagree; cf. Baron & Brown, 1991.) Yet, the complaints of skeptics about decision analysis as applied in standard "serious" domains (e.g., Ubel & Loewenstein, 1997) are actually quite consistent with the perspectives of our participants. Now suppose that the proposed interpretation is correct. Then decision analysts would be wise to make a concerted effort to convince deciders that there is indeed a strong connection between coherence and decision outcomes.

Debiasing Techniques

Debiasing techniques focus on the role of judgment in decision making. The approach developed out of a series of experimental findings demonstrating that people make systematic errors in assessing probability or, more generally, likelihood. Two of the most well-known errors are *hindsight bias* and *overconfidence*. The first is the phenomenon whereby, in hindsight, people are overly optimistic about what could have been anticipated in foresight (e.g., Fischhoff, 1975). As a concrete example, when appraising a subordinate's failed hiring decisions, a manager might be inclined to say, "*Anybody* should have been able to predict that those employees wouldn't work out." The manager would therefore be inclined to fire the subordinate for exercising poor judgment. But hindsight research suggests that the manager's opinions about what anybody would have predicted are too generous to that anybody. Thus, firing the subordinate would be ill advised. Overconfidence refers to instances – which are extremely common, in the laboratory, at least – in which people believe that their judgments are more accurate than they really are. For example, suppose that a physician makes a series of pneumonia diagnoses and, for each case, reports a probability judgment that that diagnosis is correct. Then, if the physician is overconfident, the average probability

judgment will exceed the proportion of correct diagnoses (cf. Yates, Lee, Shinotsuka, Patalano, & Sieck, 1998).

Debiasing techniques are applied in two ways. The first focuses on specific decision problems as they arise, the same way that decision analysis does. The goal is to prevent biases from injecting themselves into the judgments supporting the given, here-and-now decision. A concrete example is a set of author instructions that appears in every issue of the *International Journal of Forecasting*:

> [A]uthors of controversial papers are invited to attach a "Note to Referees." This note would describe the model (hypotheses) and possible outcomes, but *not the results*. The referees would be asked to evaluate the methodology and to predict the outcomes prior to reading the paper. They would then review the entire paper and complete the regular Referee's Rating Sheet.

Informed by research on the origins of the phenomenon, those instructions are intended to reduce the effects of hindsight bias on referees' appraisals of the manuscripts they are considering. Significantly, for the decision aid receptivity issues examined here, over the years very few authors have exercised this hindsight bias protection option (J. S. Armstrong, personal communication, May 4, 2000). The second debiasing approach entails encouraging or training deciders to alter permanently their basic judgment procedures. Their judgments would then be free of the focal biases when the decider confronts any decision problem that might present itself. Russo and Schoemaker's (1989) instructions for how managers can reduce their personal tendency to exhibit overconfidence are a good illustration (e.g., establishing a routine of seeking information that might contest one's initial opinions about any decision-related issue).

Interestingly, advocates of debiasing techniques seldom explicitly discuss why the techniques ought to be used, that is, what the advocates assume decision quality to be and why their methods should be expected to improve such quality. Perhaps that is because they consider the arguments self-evident. Such an argument could go something like the following: Most decisions are predicated at least partly on what the decider believes is likely to happen in the future, for example, that customers will respond positively to the introduction of a new product. Tendencies like hindsight bias and overconfidence amount to systematic discrepancies between deciders' judgments about what is going to occur and what really does occur. These biases impose a low ceiling

on how good the outcomes of the decisions can possibly be. Consider, for instance, a decision to introduce a new product that rests on overly optimistic predictions about how much customers will like that product. One of the few ways this decision could turn out well would be for the profit margin on each unit to be astronomically high, which is, of course, unlikely. This argument for debiasing techniques is highly compelling. It is also compatible with our participants' explanations for why they thought that a good number of their bad decisions were bad, for example, that those decisions yielded bad outcomes that they had failed to anticipate. Nevertheless, we have seen little evidence that debiasing techniques are frequently employed in actual practice. Why? Five plausible explanations suggest themselves:

- *Surprise:* It is likely that (consistent with our data), in some instances, the bad outcomes that deciders fail to predict are total surprises (e.g., side effects of a medical treatment completely new to a patient). That is, those outcomes were not even envisioned as possibilities when the decisions were being deliberated. Hence, deciders never tried to predict them in the first place. The concept of judgmental bias would be moot in such situations because there are no judgments.
- *Predictability:* It is conceivable that deciders believe that some occurrences are due entirely to chance and are, therefore, inherently unpredictable. In such a case, deciders would consider it futile to even attempt to make accurate judgments. And, once again, debiasing techniques would be seen as pointless.
- *Relevance:* Even if a decider thinks that an event is predictable to some degree, the decider might not believe that failures to anticipate the event accurately have significant implications for eventual decision outcomes. Alternatively, the decider might believe that errors in judging that event are due mainly to factors other than the biases addressed by the debiasing methods a decision consultant happens to be offering. This third possibility is implicit in some scholars' skepticism about the true practical importance of many of the biases commonly discussed in the literature (e.g., Christensen-Szalanski, 1993).
- *Efficacy:* There might well be instances in which deciders are convinced that a particular bias significantly and adversely affects judgments that, in turn, greatly increase the chances of bad decision outcomes. Yet, they may have little faith that the specific

debiasing methods being offered to them would actually work, serving to reduce the bias enough to make an investment in those methods worthwhile.

- *Habit:* Suppose that, as bias researchers contend, the judgment processes that give rise to various biases are natural. Then, even when not "hard-wired" into our basic cognitive architecture, those processes could become firmly ingrained as habits, with all the common characteristics of automaticity (cf. Anderson, 1985). In particular, those processes would tend to be evoked without conscious control or even awareness. If this is indeed so, then merely telling a decider about a debiasing technique, as is the custom, is unlikely to have any lasting impact on judgment behavior as it actually occurs in daily life.

At the moment, no one knows which, if any, of these possible explanations for the limited use of debiasing techniques are valid. But debiasing advocates would do well to find out and then revise their approaches accordingly.

Social Judgment Theory (and Its Relatives)

As the name suggests, like debiasing techniques, social judgment theory concerns itself with the judgments that typically provide essential foundations for people's decisions. The perspectives on those judgments differ markedly in the two approaches, however. Unlike debiasing methods, social judgment theory emphasizes the assumption that, in real-life judgment situations, people necessarily derive their judgments from the presumed pertinent facts or *cues* they happen to perceive in that situation. Specifically, they use their assumptions about how cues and the events of interest tend to be associated with one another in the *natural ecology*. For instance, in trying to anticipate the performance of a prospective intern, assuming that such factors tend to be predictive, a supervising physician might pay attention to the applicant's medical school record, her recommendations, and various impressions she created in her interviews. The resulting forecast would be predicated on the supervisor's beliefs about how strongly and in what form those factors generally tend to be associated with internship performance (their validity) as well as how those factors tend to covary with one another (their redundancy). Thus, all else being the same, in predicting an intern's performance, the supervisor would place heavy emphasis on factors

assumed to be strongly associated with performance and less emphasis on other, weaker (or redundant) predictors. The accuracy of the supervisor's predictions would depend on (among other things) the actual cue validities and redundancies, the correspondence between these facts and the supervisor's emphases, and how reliably the supervisor goes about the judgment task from one case to the next.

Social judgment theory traces its origins to the probabilistic functionalism espoused by Egon Brunswik and his colleagues from the 1930s to the 1950s (e.g., Tolman & Brunswik, 1935). But it has evolved over the years into a set of specific technologies and practices grounded in the spirit of Brunswik's views (cf. Brehmer & Joyce, 1988; Cooksey, 1996; Hammond, Stewart, Brehmer, & Steinmann, 1986). Social judgment theory methods provide specific means for assessing key elements of the ecology (e.g., cue validities) and of the person's judgment *policy* (e.g., the person's emphases on particular cues and his or her consistency). The theory also describes in quite precise ways how these factors lead to varying degrees of judgment accuracy. One way that social judgment theory can be "social" concerns the fact that different individuals faced with the same judgment tasks often disagree in their predictions, sometimes heatedly. Social judgment theory methods provide a means for explaining such disagreements in terms of the parties' different implicit assumptions about the ecology (e.g., about cue validities) and how they go about the judgment task (e.g., the emphases they put on particular cues and how reliably they execute their judgment policies). The isolated differences can then be examined and sometimes resolved, perhaps resulting in a collaborative judgment system that outperforms whatever existed before. Work by Cooksey, Freebody, and Davidson (1986) is illustrative. These authors derived models of the policies by which different elementary school teachers anticipated pupils' reading performance on the basis of factors such as socioeconomic status and early indices of cognitive functioning. Those models revealed significant differences in how the teachers went about this task, differences that might well have remained obscure otherwise, even to the teachers themselves. If they chose to, the teachers could then relatively easily arrive at a common consensual policy for, say, prescribing instructional interventions for students expected to have difficulty.

Important "relatives" of social judgment theory are various technologies that (a) assess the pertinent facts about a given case and then (b) use some sort of formal combination rule to synthesize those facts into a decision-relevant judgment for that case. In our internship application

example, the pertinent residency program might, for instance, have a system whereby all the major facts known about a given candidate (e.g., medical school records, referee ratings, interviewer impressions) are encoded numerically and then a performance prediction is made according to a linear equation. In that equation, the weights applied to the various encoded facts would translate into the effective impact those facts will have on the predictions. Meehl (1954) was the first to bring significant attention to the potential of such systems. Meehl showed that, under specified conditions, the systems consistently yielded more accurate clinical psychological assessments than did human diagnosticians. Later work, especially that of Robyn Dawes and his associates (cf. Dawes, 1979), provided further evidence of the efficacy of such *actuarial* judgment procedures. Just as importantly, such work used methods deriving from social judgment theory to explain *why* those systems so often outperformed humans. For instance, it showed that the most important factor is that human consistency is so grossly inferior to system consistency that it wipes out any advantages humans might have over the systems in other respects.

To the best of our knowledge, no one has done a survey to settle the issue definitively. But informal observation suggests that decision-aiding methods that draw on the ideas underlying social judgment theory and its relatives, even if not their particulars, are more commonly employed than debiasing techniques. (And if we include the formal models used in standard business forecasting for financial and marketing purposes, this is surely true.) Nevertheless, certainly in the eyes of proponents, social judgment theory and related techniques are utilized far less often than might be expected, especially given the seemingly clear-cut statistical evidence of their efficacy (see, for example, McCauley, 1991). Again, why?

To a point, whatever explains indifference to debiasing methods (including possibly the hypotheses described earlier) is likely to contribute to coolness toward social judgment theory and similar methods as well. After all, both approaches focus on the judgments people use to inform their decisions. But there are significant differences in the approaches that are likely to implicate different accounts for people's reluctance to apply them, too. For one thing, social judgment theory entails a statistical perspective that emphasizes multiple repeated instances of virtually the same situation, for example, a long history of interns and internship applicants for a residency program or a large database of mental health patients, as in the Meehl (1954) studies. In contrast, at least superficially,

for many of the biases addressed by debiasing methods (especially hindsight bias), the focus is on single cases. And the layperson deciders who participated in our studies were certainly concerned with single cases. The kinds of personally significant decision problems they discussed (e.g., what career to pursue) were in most cases one-of-a-kind problems for them as individuals, although they would not have that character for professionals seeing essentially similar cases over and over (e.g., counselors, internship admissions officials, or mental health diagnosticians). Regardless, though, the statistical-versus-single-case distinction might be significant in a fashion consistent with our data.

Recall that one thing that mattered greatly to our respondents was feeling good about the process by which they made their decisions, including the self-esteem, the pleasure, and even the sense of morality the process provided. Social judgment theory and related methods are likely to suffer on these grounds. Dawes (1979) and McCauley (1991) both convey some of the misgivings potential users have about the techniques. One is the seeming dehumanization entailed in treating all people (e.g., internship candidates) the same way, as required in a statistical approach. That approach ignores the uniqueness that is prized almost as a moral imperative in individualist cultures like that of the United States. Another is the seeming marginalization of human deciders themselves when, in the kinds of systems advocated by Dawes and Meehl, functions that were once performed by humans (e.g., clinical diagnosis) are instead performed by programmed machines. (Schoemaker & Russo, 1993, describe other interesting cases in banking.) Even worse, the machines essentially function as black boxes programmed to apply algorithms that are not rationalized in terms of the everyday causal models and language that real people prefer and use. Worse still, designed to encompass only a small number of predictive cues, the algorithms cannot accommodate the extenuating circumstances that human deciders feel compelled to take into account in individual cases (e.g., a family crisis during medical school).

The implied challenge for decision aid specialists adopting the approach of social judgment theory and related methods is to configure and present them in ways that do not ignite these negative associations. One approach would be to acknowledge explicitly people's potential misgivings and seek to dampen them (e.g., by emphasizing the fairness of treating all internship candidates the same way). Another would involve having judgments rendered by both humans and machines and then deriving a composite judgment from them, such as an average.

There are good arguments as well as evidence (e.g., McClish & Powell, 1989) that in many circumstances, such composite judgments should outperform both human and machine assessments.

General Decision Support Systems

Even among those who build, use, and study them, there is some disagreement about what exactly should and should not be considered a decision support system. Nevertheless, a definition consistent with most conceptions is that a decision support system is a computer-based system, typically interactive, that is intended to support people's normal decision-making activities (cf. Finlay, 1994; Silver, 1991). One key feature of the decision support system idea is the central role of computers. Another, which is perhaps more fundamental, is that decision support systems are not designed to alter fundamentally how the decider thinks about decision problems. And they are certainly not intended to replace the decider, making choices in his or her place. Instead, they are supposed to help (i.e., support) the decider do what he or she is inclined to do more or less naturally. Decision support systems are commonly recognized as having three basic components, configured as suggested in Figure 1.1 (cf. Carter, Murray, Walker, & Walker, 1992, p. 16):

- *Data component*: Provides substantive information the decider might request.
- *Model component*: Performs operations on information retrieved through the data component, according to models of virtually any degree of complexity.
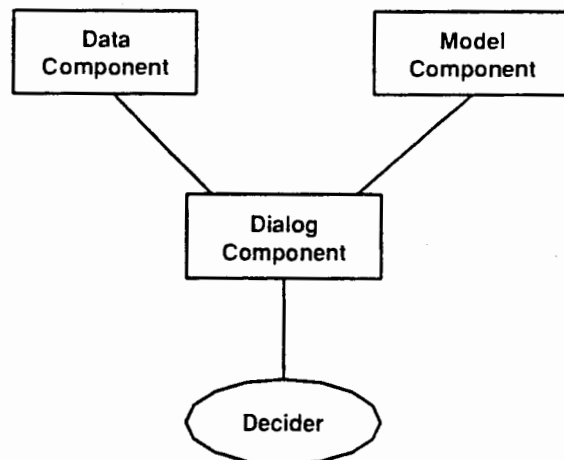

Figure 1.1 Configuration of a prototypical general decision support system.

- *Dialog component:* Allows the decider to interact with the system in particular ways, that is, constitutes the interface between the system and the decider, including things like search engines for exploring the system's databases.

We normally think of decision support systems only on the grand scale seen in large organizations such as public corporations and government agencies. But more familiar, personal-level technologies amount to decision support systems as well. Consider, for instance, an ordinary consumer (like any of us) using the on-line version of *Consumer Reports* magazine when shopping for a new refrigerator. The pertinent article on refrigerators, conveyed via the Internet, is output from the data component of the consumer's shopping decision support system, listing a host of refrigerators along with their specifications, test results, and repair statistics. The decider's personal computer, the network software, and the *Consumer Reports* displays (e.g., product × attribute tables or matrices) represent the dialog component. The typical *Consumer Reports* article provides an overall score for each of the products reviewed. These scores are a synthesis of assessments on the various features of the products – for example, temperature performance, noise – for the various refrigerators reviewed. The rules by which such scores are computed are typically weighted sums, where the weights operationalize relative feature importance. They are part of the model component of the decision support system. If the shopper wished to, he could rather easily import the data from the *Consumer Reports* refrigerator × feature matrix into his own spreadsheet. He could then perform any operations desired (e.g., compute weighted sums with importance weights different from those used by *Consumer Reports*), extending the model component of the system. It is noteworthy that the procedures commonly acknowledged as models within decision support systems include ones usually identified with decision analysis as well as social judgment theory–related tools.

A variety of aims for decision support system users are implicit in the criteria by which systems are evaluated. Generally, though, the aims fall into two broad classes, effectiveness and efficiency (cf. Finlay, 1994). *Effectiveness* refers largely to decision outcomes. Although decision support systems retain the decider's basic, natural logic, they can (and are expected to) improve outcomes because the systems' data components supply information the decider often will not have otherwise (e.g., about what refrigerators are on the market and what their specifications and

performance records are). This point does not appear to be discussed explicitly or extensively in the decision support literature. But those data components can also be expected to bring attention to objectively important considerations that naive deciders would probably overlook entirely when left to their own devices (e.g., vacuum cleaner emissions of breathable particles that can aggravate asthma and other respiratory disorders). Decision support system *efficiency* pertains to things like decision-making time and other process costs. Besides effectiveness and efficiency, decision support system developers (and, perhaps more importantly, their purchasers in businesses) pay a great deal of attention to how users feel about their experiences with the systems.

Some observers (e.g., Adelman, 1992) note that decision support systems are far less common than their early proponents anticipated that they would be (and for reasons quite similar to those argued here, such as inattention to potential users' requirements). This seems to be especially so in governmental (including military) contexts. Nevertheless, it is clear that decision support systems are much more popular than the decision aids we discussed previously. Indeed, within the business world, some kinds of systems have become virtually mainstream (see, for example, Churchill's, 1999, discussion of decision support systems for marketing). The comparative success of decision support systems is understandable, given our results. The fact that decision support systems are normally evaluated along many different dimensions is an acknowledgment of the multifaceted nature of how people conceptualize decision quality, part of our primary thesis. And the breadth of the systems (i.e., their inclusion of three distinct components that themselves are often quite broad) is very likely a response to deciders' multifaceted quality demands. It seems that commercial decision support systems would *have* to have evolved to be responsive to many of the things that deciders regard as attractive in a decision aid. Otherwise, they would have simply disappeared from the marketplace.

Similar pressures have probably ensured that the decision support systems that have survived possess certain specific advantages, too, efficiency in particular. After all, decision time and other costs are easy to assess and hence to reduce. This is especially so when there is sufficient motivation, such as the ever-present demand in the business world to control operating expenses. In their self-acknowledged pursuit of effectiveness, decision support system developers undoubtedly convey to users their aspiration to facilitate the achievement of the good outcomes our data indicate that deciders crave above all else. The data

components of decision support systems provide deciders with relevant facts that they obviously did not know before (e.g., product information from *Consumer Reports* and other databases). This surely nourishes the expectation that the systems lead to decisions with better outcomes. Whether decision support systems really do deliver all the benefits that users believe that they deliver is unclear (see, for example, Kottemann, Davis, & Remus, 1994). Yet, it *is* clear that the decision support system approach to decider receptivity has elements that the advocates of other decision aids would find worthwhile to adopt.

Group Decision Support Systems

Group decision support systems are decision support systems with one additional component beyond the three that characterize all such systems: a *group process component*. The purpose of this extra feature is to manage the interactions among several people participating in a decision-making effort. Group decision support systems acknowledge that individuals acting alone virtually never make significant decisions for organizations of any size. Instead, those decisions typically require the involvement of many people. Part of the reason is that such big decisions almost always involve political considerations. But the principle that two (or more) heads are better than one suggests that such collaboration ought to have inherent advantages for yielding objectively better decisions, too (cf. Hill, 1982). Yet, an empirically documented fact of life is that group interaction often leads to *process losses* whereby such potential goes unrealized. For instance, people sometimes withhold defensible but unconventional opinions because they fear public ridicule. Group process elements of group decision support systems seek to reduce such process losses.

In practice, group decision support tools put considerably less emphasis than other decision support systems on databases and on modeling tools and other complex information-manipulation aids. Instead, they concentrate on group process issues. This seems to be because in the kinds of decisions that are delegated to groups (e.g., corporate strategic decisions with unique circumstances and large stakes), there is often little clarity about the characteristics of the "right" decision. And, more often than not, there are fundamental differences that must be resolved concerning group members' values. Indeed, some have contended that *the* main goal of group decision support systems is to help groups work through such differences and arrive at a consensus more quickly (Olson

& Courtney, 1992). Group decision support systems usually try to do this by structuring the communication process in particular ways. For instance, they sometimes shape how input is given, when ideas are evaluated, how it is determined who speaks when, whether comments are anonymous or not, when and how voting takes place, and how individual opinions are elicited to create a climate that makes rapid consensus likely. PLEXSYS, developed by Nunamaker, Applegate, and Konsynski (1987), is a good example of a computerized group decision support tool that carries out many of the functions envisioned for such systems more generally.

Some evaluations of group decision support systems highlight group output, such as the number of new ideas generated via electronic brainstorming (e.g., Dennis & Valacich, 1993). More often, though, they tend to focus on various aspects of process, for example, how the systems change group interaction and confidence, consensus, and satisfaction of group members (cf. Dennis, George, Jessup, Nunamaker, & Vogel, 1988). Conclusions about the efficacy of group decision support systems are mixed. The systems appear to reliably yield some kinds of improvements. But they seem to pose notable challenges, too, such as increases in the social distance between group members induced by computer-mediated communications (cf. Watson, DeSanctis, & Poole, 1988). These challenges, as well as the expense, unnaturalness, and logistical difficulties associated with using computers to manage group decision making, undoubtedly help to explain why group decision support systems are far less common than more general systems (Lewis, Keleman, & Garcia, 1996). Our empirical studies did not examine participants' conceptions of group decision quality per se. Thus, our data do not speak directly to the group process aspects of group decision support systems. Nevertheless, they certainly do not disagree with anything that has been revealed in the group decision support system literature. For instance, users' negative reactions to system interference with comfortable aspects of normal, face-to-face group interaction accord well with our participants' concerns about decision process pleasantness, difficulty, and associated affect.

Expert Systems

The term *expert* in the expression *expert system* is intended to communicate either or both of two ideas. The first is that the computer program so described is supposed to perform a task, such as problem solving,

to a high degree of proficiency, at the level an expert would achieve. The second is that the system is built to mimic more or less literally the details of how a specific acknowledged human expert goes about that task. Whether modeled on any particular human expert or not, expert systems execute operations that have the qualitative nature that commonly characterizes how real people reason. For instance, they typically rely mainly on production rules of the "if-then" form, for example, "If all of the applicant's referee ratings are 4 or better, then ... " This can be contrasted to the quantitative rules that are common in the model components of most decision support systems (e.g., the linear equations of multiattribute utility theory or typical social judgment theory–related schemes). Our concern here is with expert systems that perform tasks entailed in solving decision problems. King (1990) provides an excellent illustration of a system that performs a chore normally carried out by loan officers. Specifically, the system evaluates credit requests for small business ventures, rendering recommendations to either "Give credit," "Consult a superior for more advice," or "Deny credit."

An expert system has three basic components in addition to its interface with the user (cf. Durkin, 1994; Prerau, 1990):

- *Component 1 – Knowledge base:* A repository containing essential facts, rules of thumb, or *heuristics*, and conventions about when and how to apply those facts and rules for the given domain (e.g., what information is used for appraising loan applications and what recommendations should be offered for particular configurations of facts).
- *Component 2 – Working memory:* A register for the information the user provides about the given case under consideration (e.g., storage for the particulars on Jane Smith's application for a loan to support the establishment of her software business).
- *Component 3 – Inference engine:* A processor that applies the facts, rules, and conventions in the knowledge base to the case-specific contents of working memory to perform the task set for the system (e.g., to deliver a credit appraisal for Jane Smith's new venture).

The contents of the knowledge base are critical, hence the reason that expert systems are sometimes referred to as *knowledge-based systems* or simply *knowledge systems*. Normally, the knowledge base is acquired through systematic, intensive interviews and observations of one or more recognized human experts. More generally, though, its contents

could also be collected from any source, including scholarly literature, as occurs often in medical informatics (van Bemmel & Musen, 1997). The entire process by which an expert system is built is commonly called *knowledge engineering*. Most expert systems are designed to function the way human experts behave when they serve as consultants. Conversations between experts and clients are a normal, essential element of consultations. Thus, expert systems typically contain one final feature:

- *Component 4 – Explanation module:* A set of routines that provide natural language-like explanations of how the system arrived at its conclusions for a given case (e.g., why it recommended denying a loan to Jane Smith's software company), essentially a recitation of the rules that were applied to the facts of the given case (e.g., collateral below the minimum required).

Expert systems are sometimes used as stand-alone decision aids, but there seems to be a consensus that usually they are most appropriately deployed in conjunction with decision support systems. Sometimes this amounts to simply making certain that a decider draws upon both the given decision support system and the available expert system. Alternatively (and preferably), the expert system is embedded in the larger decision support system, creating what is sometimes called a *hybrid* (Ignizio, 1991, p. 39) or *intelligent* decision support system (King, 1990, p. 53).

As in the case of decision support systems, some (e.g., Adelman, 1992) believe that expert systems have failed to live up to expectations for them or to their true potential. Nevertheless, it seems safe to say that a great many decision-related expert systems have been built and that many of them are actually in service. Durkin (1994) conducted a survey that yielded more than 600 actual expert systems for business and medical purposes, many of which we can assume to be used to assist in decision making (see also Liebowitz, 1998). Durkin estimated that his survey captured only about 20% of extant systems, which would imply more than 3,000 business and medical systems altogether. This suggests a level of popularity far beyond that of any of the decision aids we discussed previously. Why the difference?

The difference is partly due to the powerful contemporary cachet of anything that smacks of high technology. That is, businesspeople, in particular, want to at least *try* any innovation that conveys the image that their companies are on the cutting edge of the high-tech revolution

(cf. Liebowitz, 1990, p. 3). In addition, however, expert systems are developed with an eye for precisely the kinds of considerations that our data suggest are likely to drive potential users' receptivity to any decision aid. Take the case of explanation. The participants in our studies indicated that the character of the decision process matters to them a great deal. For instance, they want the process to be comfortable. Discussion with other people who might be more experienced with a given class of decision problem (i.e., experts) is routine for deciders in real life. A usual and natural feature of such discussions is give-and-take. An advisor does not simply offer a recommendation that the decider then either blindly accepts or rejects, with no questions asked. Instead, the decider requests and receives arguments for that recommendation, which the decider can then scrutinize. Unlike what occurs with the decision aids we discussed previously, this capability exists (at least minimally) for expert systems. And then there is the dominant emphasis on outcomes. Developers go to great lengths to make certain that the tasks for which they develop expert systems are manageable, well-defined ones for which expertise is clearly defined, and they avoid other, more risky tasks. Further, the human experts they choose to mimic with their systems are people who are widely accepted as experts. Accordingly, deciders almost certainly expect that using the resulting systems would yield for them the same kinds of good outcomes that presumably justified the recognition of the modeled human experts as experts in the first place.

Themes

Implicit in our review of various decision-aiding approaches are several key themes. It is useful for us to be explicit about those themes and what they imply. As Table 1.3 suggests, the six decision-aiding approaches we have discussed seem to differ sharply in their popularity. Decision analysis, debiasing techniques, social judgment theory and its relatives, and group decision support systems constitute the less popular approaches. General decision support systems and expert systems comprise the more popular class. The themes we recognize are identified with the features indicated in the last four columns of Table 1.3:

Coverage. Our main thesis has been that the typical decider recognizes several distinct facets of decision quality, that is, dimensions of goodness and ease, and that the particular aspects that are significant to any two different deciders are unlikely to coincide perfectly. To the extent

Table 1.3. *Distinctive Features of Less Popular and More Popular Decision Aid Classes*

Decision Aid Class	Feature			
	Coverage[a]	Emphasis[b]	Natural?[c]	Outcome Effects "Obvious?"[d]
Less Popular				
Decision analysis	Narrow	Procedural	Harder	No
Debiasing techniques	Narrow	Procedural	Harder (?)	No
Social judgment theory and relatives	Narrow	Procedural	Harder (?)	No
Group decision support systems	Narrow	Procedural	Harder	No
More Popular				
General decision support systems	Broad	Substantive	Easier	Yes
Expert systems	Narrow (But)	Substantive	Easier	Yes

[a] Range of quality aspects targeted.
[b] Emphasis on improving the procedure used for deciding versus providing facts about the substance of the decision problem.
[c] Whether required activities are harder or easier than customary, natural routines for deciding.
[d] Intuitive obviousness of positive effects on likely outcomes for a given decision problem.

that this is true, a decision aid's popularity is enhanced when its coverage of quality aspects is broad, that is, when it improves the decider's chances of making decisions that shine with respect to many aspects rather than only a few. As suggested by Table 1.3, all of the less popular decision-aiding approaches we reviewed have narrow coverage. For instance, decision analysis concentrates almost exclusively on ensuring the rationality of how the decider does things like synthesize probability and value (e.g., via expected utility operators). In contrast, general decision support systems tend to have very broad coverage, often actually encompassing the other aids as options along with many others among which the decider is free to pick and choose at will. Any single expert system is typically, by design, quite narrowly focused. But, because they are modular that way, several different expert systems can be offered to a decider at any one time.

Emphasis. Two broad varieties of decision aids can be recognized. *Procedural* aids seek to improve the chances of making good decisions by

improving (or replacing with better substitutes) the procedures the decider would be inclined to apply spontaneously. In contrast, *substantive* aids aim to increase the odds of achieving good outcomes by giving the decider information that is particular to the substance of the given decision problem. For instance, a substantive aid might provide the decider with previously unknown and significant facts about the options under consideration. It could even inform the decider of previously unknown options that would prove to be highly satisfactory. As indicated in the second feature column of Table 1.3, all of the less popular decision aids emphasize procedure, whereas the more popular ones highlight substance. The significance of the procedure-versus-substance distinction is likely mediated by the last two themes we discerned, which we consider next.

Naturalness. Our data indicate that deciders want their decision processes to "feel good" in several respects. For instance, they apparently feel strongly that a decision process should not be inordinately effortful or otherwise unpleasant. Thus, unless there is good reason, they would not want the process imposed by a decision aid to be more onerous than the natural ones they would apply on their own. And, ideally, the process demanded by the aid would make things easier than what is natural. Most often, the less popular decision aids we reviewed make the decision process unnatural and hard for deciders. In part, this is a necessary consequence of the fact that all those aids are procedural; their purpose is to *improve* on nature. In contrast, general decision support systems and expert systems are expressly designed to conflict as little as possible with how deciders customarily decide. Indeed, also by design, they typically make the decision process easier. For instance, via their databases, decision support systems afford ready access to decision-relevant information that deciders otherwise would find difficult, if not impossible, to acquire in a reasonable amount of time. And expert systems perform tasks that might take the decider a seeming eternity to perform, if the decider is capable of carrying out those tasks at all.

Outcome Effect Obviousness. The participants in our studies indicated that their overriding concern was that their decisions produce good outcomes for their beneficiaries. Our sense is that decision aids differ considerably in the impression that, for a given decision problem, those aids are likely to lead to a decision with better outcomes than any decision

the decider would make on his or her own. As suggested in Table 1.3, in our view (which admittedly needs to be verified for generality), for the more popular aids, general decision support systems and expert systems, the likelihood of achieving better outcomes appears obvious, whether that is actually the case or not. Such impressions seem plausible if for no other reason than that, as substantive aids, these systems routinely draw upon or provide facts that are entirely new to deciders. It is therefore easy for deciders to then say things like "My! I had no idea that was the case. That changes my mind right there." In contrast, the positive effects on decision outcomes promised by the less popular aids probably seem much more "dicey" to deciders. When it exists at all, the empirical evidence for the outcome efficacy of the new procedures entailed by these aids is statistical. When a decider arrives at a decision after applying such an aid, the decider can readily say, "For all I know, I could just as easily have arrived at the same decision doing what I normally do – and with a lot less hassle."

The Bigger Picture

The research we have described had a sharply focused aim. But as the work proceeded, we came to realize that the issues we examined, as well as our findings, have wider implications than might be immediately obvious. In this last section, we bring attention to those implications, some of which are likely to be unsettling to some observers.

Recall that an important goal of all decision scholars, and especially decision aid developers and decision consultants, is to help people decide better. However, before they can begin even trying to do that, they must convince deciders to try their wares. Earlier in this chapter, we framed this focal problem in the voice of a decider who is confronted with a decision problem and eventually concludes that she would be receptive to a suitable decision aid. It is useful for us to revisit that scenario but to extend it, to imagine that the decider does indeed choose to adopt an aid that has been offered to her, on a trial basis at least:

> Suppose I were to make this decision the way that I'm naturally inclined to make it. What degree of quality could I expect? . . . That bad, huh? It sure would be good if I had help ensuring better quality. . . . Hmmm . . . here's a decision aid that looks interesting. Perhaps *it* would do the trick. Let me look at it more carefully. . . . Seems promising. I think I'll give it a whirl.

This is the kind of soliloquy a developer or consultant would love to see, of course. It actually entails several constituent assessments. It is instructive to examine each of those assessments carefully, considering what it actually involves, how likely it is that a real decider would make that assessment (and why), and what this implies for decision aiding and for decision scholarship more generally.

Assessment 1: Decision Quality

The decider's "degree of quality" phrase implicates our primary initial focus. Our thesis was that the typical decider uses expressions like *degree of quality* to refer to a host of different things, that, subjectively, decision quality is a multifaceted construct. Further, the specific facets that any particular decider has in mind are likely to correspond only imperfectly to those that define any other decider's quality conception. As we have argued, our data agree strongly with this proposition. A variety of practical conclusions follow from the thesis, some of which were implicit in our review of major decision-aiding approaches. But one of them seems paramount and bears repeating and generalizing: The adoption of a decision aid should be enhanced significantly if that aid is constructed (and sold) as one that addresses many different aspects of decision quality, not just one.

Our data and review also highlight the significance of the decision quality concept in fundamental decision scholarship. The decision literature contains surprisingly little discussion about what the term *decision quality* ought to mean. Nevertheless, there has been *some* treatment of the subject. Authors who favor the decision analytic perspective emphasize abstract rationality, such as consistency with the axioms of utility theory or probability theory (e.g., Baron, 1988; Dawes, 1988; Edwards, Kiss, Majone, & Toda, 1984). Other authors highlight some form of *accuracy*. One form, for instance, refers to the correspondence between a decider's evaluation of an alternative and an evaluation based on a rule that some people regard as normative, such as an additive value function (e.g., Payne, Bettman, & Johnson, 1988). Another form (e.g., Frisch & Jones, 1993) emphasizes the distinction between *decision utility* and *experience utility*. The former refers to the decider's anticipatory appraisal of an alternative at the time of the decision, before it is actually selected and enacted (e.g., the appraisal of a book being contemplated for purchase). The latter describes the decider's assessment of the actual experience with an alternative (e.g., the decider's degree of liking

or disliking a book the decider actually reads). The pertinent decision quality conception emphasizes the correspondence between decision and experience utility; the higher the correspondence, the better. Yet another class of quality conceptions focuses on the decision process, highlighting process features that arguably ought to be expected to enhance the decider's satisfaction with chosen alternatives (e.g., Frisch & Clemen, 1994; Janis & Mann, 1977). Each of these definitions of decision quality can be (and has been) argued to be inadequate in some way. And there is clearly no universal agreement on which definition should be accepted.

Consensus on a suitable definition would be good for decision scholarship generally and certainly for decision aiding. After all, if we are uncertain about where we are trying to go (e.g., helping people achieve good decisions or understanding the nature of what deciders themselves consider to be decision making), how can we tell when we have arrived there? Moreover, consensus would probably allow for more productive and efficient scholarly discussions and collaborations. We hold no illusions that definitional consensus will occur any time soon. Achieving consensus on a decision quality conception that is not vacuously broad but nevertheless accommodates the multifaceted character implicit in everyday practical and scholarly usage is a formidable task. Nevertheless, we offer for initial consideration a set of quality definitions that we have employed for some time in our own work (cf. Yates & Estin, 1998; Yates & Patalano, 1999). These definitions appear to serve a useful focusing function. Importantly, they are also consistent with the responses of participants in the studies reported here and with decision-making practices in a variety of practical domains we have examined (e.g., business, medicine, personal counseling, aviation). The definitions:

Good Decision. A good decision is one that is strong with respect to one or more of the following five criteria:

1. *The aim criterion:* The decision meets the decider's explicitly formulated aim or aims (e.g., when a company sets out to hire a person who keeps good financial records and the chosen applicant does indeed keep good records).
2. *The need criterion:* The decision satisfies the actual needs of the beneficiary, needs that may or may not correspond to the decider's aim(s) (e.g., when a company needs someone to maintain

good financial controls, and in fact hires such a person, even if by accident, as when the company merely searched for a person who keeps good records).

3. *The aggregated outcomes criterion:* Collectively, all of the actual outcomes of the decision, including ones beyond particular aims and needs, are better than the active reference, such as the status quo or the beneficiary's aspiration level (e.g., when a company hires a candidate who keeps good records, meets its cost control needs, and moreover solves the company's morale problems, thereby leaving the company much better off than before).

4. *The rival options criterion:* In the aggregate, the outcomes of the decision are superior to those that would have resulted from any and all available competing alternatives (e.g., when, taking everything into account, a company is better off having hired Person X than it would have been hiring any other candidate on the market).

5. *The process costs criterion:* The costs of arriving at the decision are minimal (e.g., in money, time, effort, or aggravation) (e.g., when the process of searching for and reviewing job candidates not only causes little interference with normal work routines but is actually pleasant for all involved).

Good Decision Process. A good decision process is one that tends to yield good decisions.

With its emphasis on outcomes, our definition of a good decision differs sharply from the one that holds sway among decision analysts, and intentionally so. Those who practice decision analysis often point-edly seek to discourage deciders from their natural inclination to ap-praise decisions according to outcomes. For example, some urge de-ciders to recognize that "in an uncertain world where unforeseeable events are common, good decisions can sometimes lead to bad out-comes" (Hammond, Keeney, & Raiffa, 1998, p. 52). The present quality conceptions would discourage such statements; "good decision" and "bad outcomes" are inherently contradictory in that view, as our data suggest they are to the typical decider. The proposed notion of a good decision process, with its focus on what the process *tends* to produce, is statistical. Thus, it acknowledges the inescapable fact of sampling that not every decision made by a good real-life decision process will result in good outcomes (i.e., strength with respect to every quality criterion).

That conception therefore leads to the endorsement of statements like this: "Decisions that are made by a good *process* can sometimes lead to bad outcomes."

It is important to recall that, from the decision analytic perspective, a good decision is defined to be one that adheres to principles of rationality, such as expected utility maximization, which rest on notions of logical consistency or coherence. Our data suggest that most deciders would be content with pursuing this kind of decision goodness only if there were compelling evidence that doing so would, even if only on average, be rewarded by superior outcomes. As we noted previously, such evidence does not appear to have been reported. Instead, advocates of coherence conceptions of decision quality have relied on the plausibility of a strong coherence–outcomes link. In contrast, besides being statistical, the proposed alternative conception of decision quality is patently empirical. Decisions made according to a good process should be documented as typically yielding few serious shortcomings in terms of the various quality criteria, whether this is the result of adherence to principles of coherence or anything else.

Investigators have sometimes described to two different groups of research participants the circumstances surrounding some decision problem (e.g., a medical treatment dilemma) and the decision that was made in the pertinent situation. One group is then told that the decision turned out well (e.g., the patient recovered), but the other is told that the outcome was bad (e.g., the patient died). When asked for their opinion of the decision, participants in the former group tend to characterize that decision as better than do those in the latter group. Investigators who hold to the decision analytic conception of decision quality (e.g., Baron & Hershey, 1988) consider such results to be evidence of an error called *outcome bias*. In the convention proposed here, they are not. Rather, they are merely a reflection of people's customary semantic preferences. Studies of outcome bias have revealed useful facts about such preferences. Yet, in our view, such research could be even more revealing of fundamental principles if we simply acceded to those preferences, as in the proposed decision quality conception.

Assessment 2: Need

The second assessment made by the decision aid–receptive decider in our scenario is that she would welcome – she needs – a decision aid because she would expect her decision to have poor quality if she were

left to her own devices. *("That bad, huh?")* This perception of need is clearly critical. After all, if there is no need, why bother with the very idea of a decision aid? For the moment, let us suppose that when she speaks of decision quality, our decider is referring to decision products rather than the decision process, for example, the costs of deciding. The uncomfortable conclusion suggested by our data as well as other evidence is this: People like the decider in our scenario are rare. Unlike her, typical deciders are likely to believe – rightly or wrongly – that they are perfectly capable of making most decisions well, with no assistance whatsoever. If true, this conclusion would, in and of itself, constitute an exceptionally compelling explanation for across-the-board indifference to decision aiding and a formidable hurdle for any decision aid developer or consultant to overcome.

So, what is the evidence? Recall that, in Study 1, participants were more satisfied with the outcomes of their easy decisions than those yielded by their hard decisions. Yet, in absolute terms, on average participants were highly pleased with how *all* their decisions turned out. Also recall that, in Study 2, participants regarded the badness and importance of their bad decisions as less extreme than the goodness and importance of their good decisions. Moreover, they were faster in bringing to mind their worst decisions than their best ones, plausibly because those decisions were especially distinctive in their unusualness. These data agree with the proposition that people generally think that their unaided decision making in the past was, for the most part, just fine. So why should they not expect that their future decisions will be just as good?

Results from a variety of other studies also suggest that people have high expectations that their decisions will turn out well. Those results further suggest that these expectations might well be too high, and for specific reasons. First are findings in the literature on *bolstering* (cf. Janis & Mann, 1977). This is the oft-documented phenomenon whereby, after a person has chosen some object (e.g., a work of art), the person's appraisal of that object, relative to the rejected competitors, becomes more favorable. Theories used to explain bolstering cite the tendency for deciders to enhance their perceptions of the attractiveness of those aspects of the chosen object that happen to be strong rather than weak. Thus, the possibility of bolstering should be facilitated for decisions involving ambiguous and multifaceted consequences as compared to those entailing simpler ones (e.g., complex paintings vs. sacks of flour). As our data have demonstrated, people are inclined to appraise typical

real-life decisions with respect to several aspects rather than just one. This therefore implies an especially high likelihood that people will remember their decisions as good ones, perhaps even better than they had anticipated in the first place.

Another literature suggests that, even when a decision yields outcomes that are both objectively and subjectively bad at the moment they occur (e.g., disability from an accident), the decider will not indefinitely experience and remember them that way (cf. Kahneman, 1999). Instead, the experience will be effectively neutralized, plausibly through mechanisms such as the development of lower adaptation and aspiration levels (e.g., lower mobility goals than before the disability). The pertinent literature also invites a related and especially powerful speculation. It suggests that, no matter what their circumstances might be, people are highly likely to reason as follows: "I'm quite content with the way things are. So why do I need to make any radical decisions of any kind? And if I don't need to make any big decisions, then I certainly don't need any *help* with those decisions."

Attribution processes are yet another reason to expect deciders to recall their past decisions as good ones, no matter what. There is credible documentation of the tendency for people to attribute to their own actions (e.g., their own decisions) the good things that happen to them (cf. Ross & Fletcher, 1985), but when bad things occur, they are inclined to attribute them to other factors, such as chance, essentially explaining them away. Dispositions like these plausibly have surprising, far-reaching implications. For instance, even when people or organizations are in dire straits because of bad prior decisions, they are unlikely to acknowledge (or even recognize) that fact. Instead, they are inclined to say things like "Well, yes, we're in a bit of a bind here, but it's just a run of bad luck. It has nothing to do with how we make decisions. So there's no point in us even talking about things like decision aiding." Evidence for tendencies like these has not been limited to people's appraisals of their past decisions. It is probably manifest in reports of particular kinds of overconfidence, too. A good example is the tendency for trial lawyers to be overly optimistic that they will win the cases they choose to take to court (e.g., Loftus & Wagenaar, 1988).

Now suppose that people do, in fact, seldom perceive a need for help in making decisions that will turn out well for them. What should a decision aid developer or consultant do? Seek another vocation because the demand for his or her services simply does not exist? Several less extreme and more defensible alternatives suggest themselves.

Option 1 – Emphasize Ease. The first alternative is to abandon the aim of helping people make decisions with better outcomes and instead develop aids that help people make the same decisions they would make anyway, but more easily. As we saw before, this is effectively the goal that many decision support systems and expert systems pursue and achieve. And, as we also conjectured, prospective decision-aiding clients readily see the value of such assistance and hence welcome it.

Option 2 – Destroy Illusions. A second alternative is predicated on the assumption that people's belief in the adequacy of their own past and future decisions is at least partly illusory. That is, deciders could benefit from assistance but do not realize that fact. The tack that a decision aid developer or consultant could then take would be to enlighten the decider, to destroy the illusions. It seems that, if the opportunity exists, an especially effective way to do this is to have the decider keep accurate *decision logs.* That is, before major decisions are made, the decider would record things like the aims of those decisions. And then, immediately after the decisions are made and their outcomes are determined, the decider would record what those outcomes were. Periodically, those records would be reviewed. To the extent that the decider's quality appraisals are indeed illusory, outcomes should systematically fall short of expectations. And the decider should be convinced of the actual need for assistance. A derivative of this second approach would be especially appropriate, it seems, in organizations. Although the various mechanisms we have reviewed should serve to reinforce people's illusions about the efficacy of their own decisions, they should not influence appraisals by other people. That is, although we ourselves might think that our decisions are consistently outstanding, others would not evaluate them so generously. Thus, a good way to deflate self-appraisals would be to complement them with appraisals by others. In the typical organization, it is those other appraisals that matter the most anyway. Of course, in business organizations, financial outcomes are unambiguous and typically have central significance. Nevertheless, even there, immediately assessable financial outcomes of decisions are almost never the sole concern.

An important prerequisite of this second option should not go unnoticed: The decider must be receptive to being exposed to arguments and evidence that his or her decisions are at greater risk than suspected. Our hunch is that this is an especially significant hurdle for decisions that any given person makes only a few times in life, such as career

choices, home purchases, and the decision about whether to marry a current suitor. And because of emotional factors and a perception of the uniqueness of one's personal circumstances, arguments about the risks of marriage decisions are especially prone to fall on deaf ears, despite widespread awareness that the divorce rate is extremely high in many places, nearly 50% in the United States.

Option 3 – Frame Bad Decisions as Pathologies and Traumas. A final approach to the problem of deciders perceiving no need for assistance is radical. Even taking into account some measure of self-serving illusion in people's appraisals of their own decisions, this approach concedes the possibility that, on average, most people's decisions *do*, in fact, yield good outcomes for them. This view rests on an analogy with health care. The vast majority of the population is in decent if not good physical health. On the other hand, everyone occasionally gets sick, and some experience traumas of various kinds, such as injuries from car accidents. The health care system, which is clearly indispensable, is designed to deal with these unusual occurrences. The public health segment of the system seeks to reduce their incidence, whereas the medical segment treats them when they happen to occur. This *pathology/trauma model* could serve as a prototype for how to design a *decision care system*. We all realize that, although we might be in perfect health today, that situation could change dramatically in an instant. Part of the mission of a decision care system would be to convince people that they are at risk for rare but catastrophic decisions, too. Physical illness and trauma are obvious. A challenge for decision specialists is to make catastrophic decision failures obvious, too. A second challenge would be to design decision aids that perform preventive functions analogous to public health interventions. A third would be to create aids that yield decisions that minimize the damage when outcomes are bad (i.e., mitigate risks) or that allow for rapid recovery from failed decisions.

Assessment 3: Relevance and Efficacy

This was the third important assessment made by the decider in our scenario: *"Seems promising (as a means of ensuring better decision quality)."* The decider has actually made two distinct constituent assessments within this larger one. The first is that the focal decision aid is relevant to the decision problem at hand. That is, it addresses factors that, in turn, affect decision quality facets the decider cares about. The focus of

our empirical studies about hard-versus-easy and good-versus-bad decisions was on this kind of assessment, identifying personally relevant quality dimensions. The second assessment is that the aid in question would actually work, that it really *would* improve the chances of making decisions that are good with respect to the particular aspects of decision quality deemed relevant by the decider. Our sense is that, as implicit in our review of major decision-aiding approaches, this kind of assessment must be critical to people's receptivity to a decision aid. After all, if a decider thinks that an aid addresses a personally significant aspect of decision quality (e.g., exposure to risk) but has little chance of actually delivering good quality, why bother with it?

Underneath the decider's judgment about an aid's odds of facilitating good decisions must be a personal *naive decision attribution theory* of the form illustrated in Figure 1.2(a). That theory is the decider's conception of how the process by which he or she decides yields a decision that, in turn, achieves or falls short of the decider's multifaceted standard of good decision quality. The theory also includes beliefs about the nature of any decision aid that is available. When pondering whether to adopt that aid, the decider speculates about how, if applied, it would affect various aspects of decision quality, both directly and indirectly through its effects on elements of the decider's process. For instance, suppose the decider believes that when decisions in the relevant class fail, it is mainly because the decider is ignorant of critical facts. Then the decider will adopt an aid only if it appears that it would supply those facts.

To the best of our knowledge, no one has actually studied naive decision attribution theories. But we speculate that these theories have interesting and important properties. It seems doubtful that, according to these theories, a significant contributor to failed decisions is that deciders synthesize facts in ways that disagree with, say, expected utility or multiattribute utility maximization. To the extent that this is true, this implicates a contributor to people's indifference to forms of decision analysis that emphasize such rules. Parallel to deciders' naive decision attribution theories are *valid decision attribution theories*, as sketched in Figure 1.2(b). Rather than deciders' speculations, these theories describe actual process and decision aid features and their true influences on decisions and decision quality. Thus, although deciders themselves might think that, say, expected utility or multiattribute utility rules are irrelevant to decision quality, that may or may not be the case in reality. An important part of the future agenda for decision research is the accurate depiction of both naive and valid decision attribution theories. In the
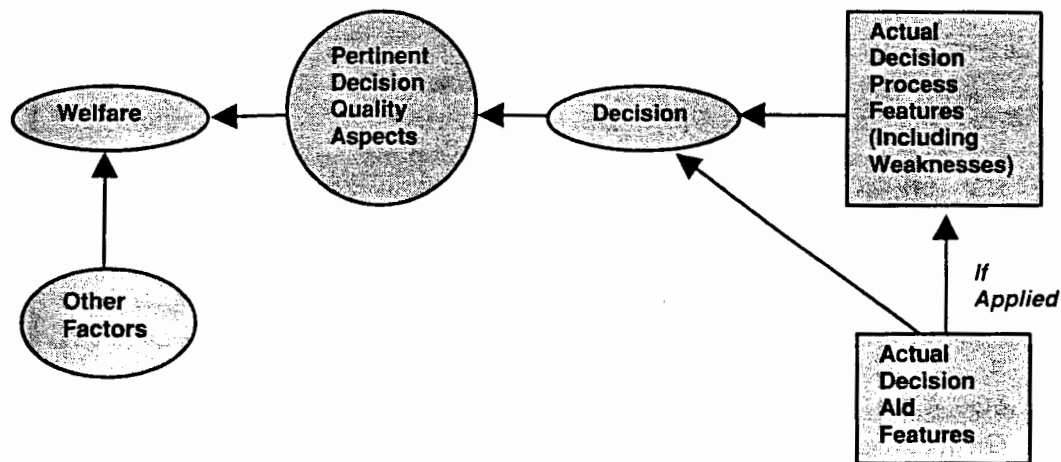
a)
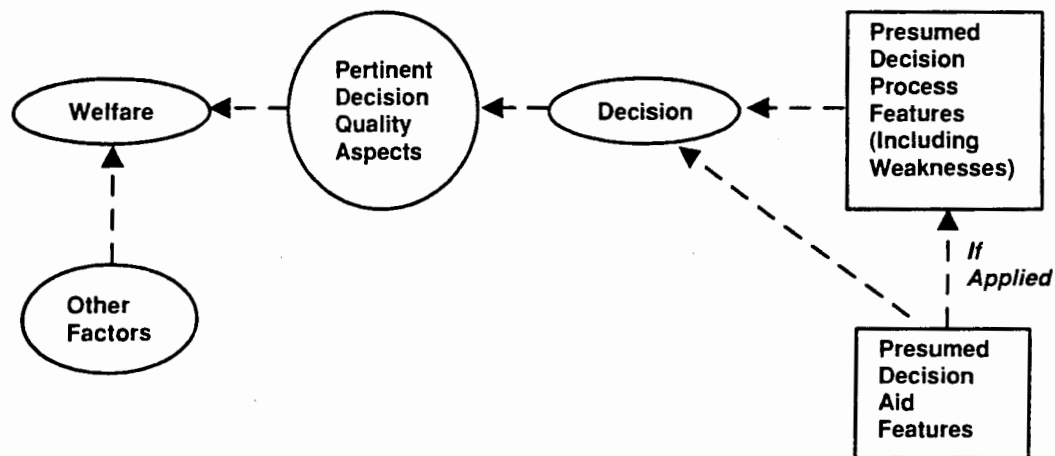
b)

Figure 1.2 Schematic representation of decision attribute theories: (a) naive and (b) valid.

discussion thus far, we have limited ourselves to attributions for decision quality. We predict, though, that the most useful theories will be ones that go a step further and, as suggested in Figure 1.2, encompass the decision beneficiary's welfare as well. For instance, the data and ideas reviewed in this chapter suggest that it is not at all unlikely that many people believe that their personal welfare – what really matters to them – is largely a function of factors that have nothing to do with their

decisions. People holding such beliefs would, understandably, have little interest in decision aids or, for that matter, decision making.

References

Aaron, M. C. (1995). The value of decision analysis in mediation practice. *Negotiation Journal*, *11*(2), 123–133.

Adelman, L. (1992). *Evaluating decision support and expert systems.* New York: Wiley.

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.) New York: Freeman.

Aoki, N., Kitihara, T., Fukui, T., Beck, J. R., Soma, K., Yamamoto, W., Kamae, I., & Ohwada, T. (1998). Management of unruptured intracranial aneurysm in Japan: A Markovian decision analysis with utility measurements based on the Glasgow outcome scale. *Medical Decision Making*, *18*, 357–364.

Baron, J. (1988). *Thinking and deciding.* New York: Cambridge University Press.

Baron, J., & Brown, R. V. (Eds.). (1991). *Teaching decision making to adolescents.* Hillsdale, NJ: Erlbaum.

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*, 569–579.

Brehmer, B., & Joyce, C. R. B. (Eds.). (1988). *Human judgment: The SJT view.* Amsterdam: North-Holland.

Carter, G. M., Murray, M. P., Walker, R. G., & Walker, W. E. (1992). *Building organizational decision support systems.* Boston: Academic Press.

Christensen-Szalanski, J. J. J. (1993). A comment on applying experimental findings of cognitive biases to naturalistic environments. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 252–261). Norwood, NJ: Ablex.

Churchill, G. A., Jr. (1999). *Marketing research: Methodological foundations* (7th ed.). Fort Worth: Dryden.

Clemen, R. T. (1991). *Making hard decisions: An introduction to decision analysis.* Boston: PWS-Kent.

Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications.* San Diego: Academic Press.

Cooksey, R. W., Freebody, P., & Davidson, G. (1986). Teachers' predictions of children's early reading achievement: An application of social judgment theory. *American Educational Research Journal*, *23*, 41–46.

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

Dawes, R. M. (1988). *Rational choice in an uncertain world.* San Diego: Harcourt Brace Jovanovich.

Dennis, A., George, J., Jessup, L., Nunamaker, J., & Vogel, D. (1988). Information technology to support electronic meetings. *MIS Quarterly*, *12*, 591–624.

Dennis, A. R., & Valacich, J. S. (1993). Computer brainstorms: More heads are better than one. *Journal of Applied Psychology*, *78*, 531–537.

Durkin, J. (1994). *Expert systems design and development.* New York: Macmillan.

Edwards, W., Kiss, I., Majone, G., & Toda, M. (1984). What constitutes "a good decision?" *Acta Psychologica, 56,* 5–27.

Finlay, P. (1994). *Introducing decision support systems.* Oxford: Blackwell.

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.

Frisch, D., & Clemen, R. T. (1994). Beyond expected utility: Rethinking behavioral decision research. *Psychological Bulletin, 116,* 46–54.

Frisch, D., & Jones, S. K. (1993). Assessing the accuracy of decisions. *Theory and Psychology, 3,* 115–135.

Hammond, J. S., Keeney, R. L., & Raiffa, H. (1998). The hidden traps in decision making. *Harvard Business Review, 16*(5), 47–58.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1986). Social judgment theory. In H. R. Arkes & K. R. Hammond (Eds.), *Judgment and decision making* (pp. 56–76). Cambridge: Cambridge University Press.

Hax, A. C., & Wiig, K. M. (1977). The use of decision analysis in a capital investment problem. In D. E. Bell, R. L. Keeney, & H. Raiffa (Eds.), *Conflicting objectives in decisions* (pp. 277–297). New York: Wiley.

Hill, G. W. (1982). Group versus individual performance: Are $N + 1$ heads better than one? *Psychological Bulletin, 91,* 517–539.

Ignizio, J. P. (1991). *An introduction to expert systems: The development and implementation of rule-based expert systems.* New York: McGraw-Hill.

Janis, I., & Mann, L. (1977). *Decision making.* New York: Free Press.

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York: Russell Sage Foundation.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291.

Keeney, R. L., & Nair, K. (1977). Selecting nuclear power plant sites in the Pacific Northwest using decision analysis. In D. E. Bell, R. L. Keeney, & H. Raiffa (Eds.), *Conflicting objectives in decisions* (pp. 298–322). New York: Wiley.

King, D. (1990). Modeling and reasoning: Integrating decision support with expert systems. In J. Liebowitz (Ed.), *Expert systems for business and management* (pp. 51–76). Englewood Cliffs, NJ: Yourdon Press.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). *Decision making in action: Models and methods.* Norwood, NJ: Ablex.

Kottemann, J. E., Davis, F. D., & Remus, W. E. (1994). Computer-assisted decision making: Performance, beliefs, and the illusion of control. *Organizational Behavior and Human Decision Processes, 57,* 26–37.

Lewis, L. F., Keleman, K. S., & Garcia, J. E. (1996). Possible barriers and challenges to the adoption of group support systems. *Group Decision and Negotiation, 6,* 189–194.

Liebowitz, J. (Ed.). (1990). *Expert systems for business and management.* Englewood Cliffs, NJ: Yourdon Press.

Liebowitz, J. (Ed.). (1998). *The handbook of applied expert systems.* Boca Raton, FL: CRC Press.

Loftus, E. F., & Wagenaar, W. A. (1988). Lawyers' predictions of success. *Jurimetrics Journal, 16*, 437–453.

Matheson, J. E., & Howard, R. A. (1968). An introduction to decision analysis. In Decision Analysis Group (Eds.), *Readings in decision analysis* (pp. 7–43). Menlo Park, CA: Stanford Research Institute.

McCauley, C. (1991). Selection of National Science Foundation graduate fellows: A case study of psychologists failing to apply what they know about decision making. *American Psychologist, 46*, 1287–1291.

McClish, D. K., & Powell, S. H. (1989). How well can physicians estimate mortality in a medical intensive care unit? *Medical Decision Making, 9*, 125–132.

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

Morris, W. T. (1977). *Decision analysis.* Columbus, OH: Grid, Inc.

Nunamaker, J. F., Applegate, L. M., & Konsynski, B. R. (1987). Facilitating group creativity with GDSS. *Journal of Management Information Sciences, 3*(4), 5–19.

O'Connor, A. M., Rostom, A., Fiset, V., Tetroe, J., Entwistle, V., Llewellyn-Thomas, H., Holmes-Rovner, M., Barry, M., & Jones, J. (1999). Decision aids for patients facing health treatment or screening decisions: Systematic review. *British Medical Journal, 319*, 731–734.

Olson, D. L., & Courtney, J. F. (1992). *Decision support models and expert systems.* New York: Macmillan.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*, 534–552.

Prerau, D. S. (1990). *Developing and managing expert systems: Proven techniques for business and industry.* Reading, MA: Addison-Wesley.

Ross, M., & Fletcher, G. J. O. (1985). Attribution and social perception. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 73–122). New York: Random House.

Russo, J. E., & Schoemaker, P. J. H. (1989). *Decision traps: The ten barriers to brilliant decision-making and how to overcome them.* New York: Doubleday.

Schoemaker, P., & Russo, J. E. (1993). A pyramid of decision approaches. *California Management Review, 36* (1), 9–31.

Silver, M. S. (1991). *Systems that support decision makers: Description and analysis.* New York: Wiley.

Tolman, E., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review, 42*, 43–77.

Ubel, P. A., & Loewenstein, G. (1997). The role of decision analysis in informed consent: Choosing between intuition and systematicity. *Social Science and Medicine, 44*, 647–656.

van Bemmel, J. H., & Musen, M. A. (Eds.). (1997). *Handbook of medical informatics.* Houten, the Netherlands: Springer.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* New York: Cambridge University Press.

Watson, R. T., DeSanctis, G., & Poole, M. S. (1988). Using a GDSS to facilitate group consensus: Some intended and unintended consequences. *MIS Quarterly, 12*, 463–478.

Yates, J. F. (1990). *Judgment and decision making*. Englewood Cliffs, NJ: Prentice Hall.

Yates, J. F., & Estin, P. A. (1998). Decision making. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 186–196). Malden, MA: Blackwell.

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

Yates, J. F., & Patalano, A. L. (1999). Decision making and aging. In D. C. Park, R. W. Morrell, & K. Shifren (Eds.), *Processing of medical information in aging patients: Cognitive and human factors perspectives* (pp. 31–54). Mahwah, NJ: Erlbaum.

Zsambok, C. E., & Klein, G. (Eds.). (1997). *Naturalistic decision making*. Mahwah, NJ: Erlbaum.