

The Real World: What Good Is It?

BARUCH FISCHHOFF

Carnegie Mellon University

Any experimental science runs the risk of implosion. The tasks that were created as surrogates for complicated real-world situations can take on lives of their own. When that happens, scientists become fascinated with the nuances of variations within that little world. Their theoretical accounts end up with little place for phenomena that could not be observed there. Extrapolations to other settings may require large doses of conjecture (or leaps of faith). This essay examines some possibilities for understanding the world of the lab by grappling with applied problems in the world outside and perhaps doing a little good along the way.

© 1996 Academic Press, Inc.

Choice A

In this task, you will be asked to choose between a certain loss and a gamble that exposes you to some chance of loss. Specifically, you must choose either: Situation A. One chance in 4 to lose \$200 (and 3 chances in 4 to lose nothing). OR Situation B. A certain loss of \$50. Of course, you'd probably prefer not to be in either of these situations, but, if forced to either play the gamble (A) or accept the certain loss (B), which would you prefer to do? (Fischhoff, Slovic, & Lichtenstein, 1980)

Choice B

My cousins . . . ordinarily, I'm like really close with my cousins and everything. My cousin was having this big graduation party, but my friend—she used to live here and we went to . . . like started preschool together, you know. And then in 7th grade her

stepdad got a job in Ohio, so she had to move there. So she was in Ohio and she invited me up for a weekend. And I've always had so much fun when I'd go up there for a weekend. But, it was like my cousin's graduation party was then, too—like on that same weekend. And I was just like I wanted to go to like both things so bad, you know. I think I wanted to go more to like up to Ohio, you know, to have this great time and everything, but I knew my cousin—I mean, it would be kind of rude to say, "Well, my friend invited me up, you know for the weekend." And my cousins from out of town were coming in and everything. So I didn't know what to do. And I wanted my mom to say, "Well, you have to stay home," so then I wouldn't have to make the decision. But she said, "I'm not going to tell you, you have to stay home. You decide what you want to do." And I hate when she does that because it's just so much easier if she just tells you what you have to do. So I decided to stay home basically because I would feel really stupid and rude telling my cousin, well, I'm not going to be there. And I did have a really good time at her graduation party, but I was kind of thinking I could be in Ohio right now. (Fischhoff, Furby, Quadrel, & Richardson, 1991)

The first of these choices looks like the kind of tidy stimulus found in many studies of behavioral decision making. The second looks like the kind of messy process found in many people's minds, when they try to grapple with what life sends their way. Our goal in studying choices like A is to illuminate situations like B. A nagging worry has to be whether you can get there from here.

There are good reasons for studying A, rather than B. So much is going on in B that it is hard to tell which factors influence behavior. Choice B is so unique that one cannot say how people generally behave in such situations, at least without grappling with what "such situations" means. The richness of Choice B comes through the filter of verbal reports, which can add both random and systematic error (e.g., Ericsson & Simon, 1984; Nisbett & Wilson, 1977). Observing Choice B *in vivo*, and even eliciting concurrent verbal protocols, might reduce these problems at the price of incurring others (e.g., influencing decision makers' behavior).

Other professions (e.g., history, ethnography) have complementary reasons for avoiding the risks that come with the standardization and experimentation

Talks developing the ideas in this article were given as colloquia at Princeton, Rice, Carnegie Mellon, and New York Universities, and as a Presidential Address to the Society for Judgment and Decision Making. The comments of those audiences are gratefully appreciated, as are the contributions of the co-authors and investigators cited here. Shane Frederick, Gideon Keren, Ola Svenson, and an anonymous reviewer provided valuable comments on a draft of this manuscript. Its preparation was supported by the National Institute on Alcohol Abuse and Alcoholism and the National Science Foundation. The opinions expressed are, of course, those of the author. Address reprint requests to Baruch Fischhoff, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213.

that we prefer. Some are so comfortable with studying the particular that they deny the wisdom (or even possibility) of generalization (e.g., deriving laws of history) (e.g., Gallie, 1964; Hexter, 1971). They might be skeptical of our hope of eventually informing specific decisions with general insights (Kren & Rapoport, 1976). Even psychologists committed to generalization have cautioned against the premature closure that can come with neglecting detail (Frisch, 1993; Gergen, 1994; Gilligan, 1982)

This essay considers some ways to bridge the gap between the general and the specific in decision making, driven by the demands of trying to be useful in real-world settings. Responses to these demands have been divided into *applied basic* psychology and *basic applied* psychology (Baddeley, 1979). The former uses basic research tools to solve applied problems, possibly learning something general from the obstacles encountered in the process. The latter looks for unsolved basic issues embedded in applied problems. The former may be more generous (giving psychology away), while the latter may be more modest (acknowledging our need to learn). Both can provide our science with exogenous sources of change, complementing the endogenous ones that arise from self-generated problems (Kuhn, 1962; Lakatos, 1970).

The set of studies described below begins and ends with attempts to look simultaneously at many aspects of the same decision. In the middle are studies focused on components of decisions (options, consequences, uncertainties). In each case, the research attempts to learn something general by wrestling with the particulars of specific decisions. A recurrent concern in these studies is whether investigators' conceptualizations of decisions match those of the people making them. At the level of individual studies, identifying discrepancies can clarify the interpretation of results. Looking across studies, the pattern of discrepancies (and convergences) might contribute to a theory of context, showing how people impose structure on the choices that life presents them.

These studies are chosen from my own work, conducted with several valued colleagues. This restriction allows me to speak from personal knowledge regarding the practical concerns that motivated each study, as well as how we tried to "make science" out of the issues that it raised. I try to be candid about how far these studies have made either basic or applied contributions. The exposition presents enough results to give some feeling for whether this level of detail contains a message or just a mess. Some related efforts by other investigators are noted, although space constraints prohibit doing them justice.

WHOLE DECISION STRATEGIES I: TEEN DECISIONS

Some years ago, we received support, from the Carnegie Council on Adolescent Development, to study adolescent risk taking. Part of our response was to replicate, with teens, structured experiments that had previously been conducted with adults, in which subjects judge the magnitude of risks and the extent of their knowledge about them (Quadrel, Fischhoff, & Davis, 1993; Quadrel, Fischhoff, Fischhoff, & Halpern, 1994a). Teens' responses were similar to those observed with adults, a mixed pattern of strengths and weaknesses. Indeed, they were so similar as to raise the suspicion that the method had shaped the message. Especially given the deep concern over adolescents' "reckless" behavior (e.g., Arnett, 1993; Elkind, 1967), we worried that our methods had prompted "adult-like" modes of thought, unlike adolescents' normal way of thinking, out there in the world.

As a result, we supplemented these structured studies with more open-ended ones, allowing a wider range of thought processes to emerge.¹ Fortunately, our funding carried relatively weak strings, allowing us to attempt much more speculative studies than were likely to have survived ordinary proposal review.

The most ambitious of these studies involved 1- to 2-h interviews with 105 teenage girls, drawn from disparate social backgrounds, in Eugene and Pittsburgh. After a brief introduction, these teens were asked to describe briefly recent decisions in several specific domains (e.g., about school, parents, clothes). Then, they described more thoroughly three difficult decisions of their own choosing, two made recently and one still facing them. They used their own words, shaped only by our prompts to consider specific issues (after their spontaneous descriptions had ebbed). Choice B is the initial description of one of these decisions. Our coding of the transcripts considered the content, structure, and process of these decisions. What follows is a sampling of results from each of these three areas. These results are from 66 younger teens (age: 12–14; grade: 6–8) from varied socioeconomic backgrounds, with the median being lower-middle class. Fischhoff *et al.* (1991) provide a full report.

¹ Of course, there is more to the behavior of adolescents, or others, than cognition (Fischhoff, 1992). Although our studies allowed teens to discuss issues like emotions and social pressure, it is still just talk and far removed from observing real-life decision making in its full complexity.

Content

Those who study decision making in general often have relatively little interest in the topics of everyday decisions. Their primary concern is processes that emerge across domains. It would be a diversion for them to develop a theory of context, characterizing decision topics and examining their effects on thought processes.² One aspect of content that sometimes does appear in theoretical discussions is familiarity. Performance may improve if familiarity provides fluency with the issues, making it easier to focus on task structure. Or, familiarity may reduce performance, if it encourages habitual responses, leading people to ignore task structure. Familiarity may also affect the vocabulary for discussing a domain, perhaps stabilizing usage, perhaps leading to the divergence found in a living language. Unrecognized differences in usage can produce tasks that mislead subjects and responses that mislead investigators. People should have better elaborated beliefs about topics that are on their minds. They may also have more accurate beliefs, if they have had the opportunity to learn. Thus, how people speak can illuminate the experiences that they have had.

In order to find out what topics were on these teens' minds, we asked them to describe briefly recent decisions in each of seven domains (school, free time, clothing, friendships, health, money, and parents). They had no trouble doing so, producing 17.4 decisions per subject. The largest category of these decisions dealt with friendships (4.6), while very few concerned the sort of fateful issues that concern parents and educators (e.g., drugs, sex, violence, alcohol). Such "risk behaviors" were, however, somewhat more common when the teens later described three hard decisions in detail. The most common topic of hard decisions was openness with friends, followed, at some distance, by openness with parents, drugs, career plans, and school performance.

Of course, these domains are not independent. For example, relations with friends, parents, and teachers shape and are shaped by decisions about drugs. Our subjects' reports suggest that health-risk issues emerge sequentially, after other decisions have been made. Experimental research has demonstrated predictable differences in people's choices when decisions are framed in simultaneous or sequential form; moreover, people are

unlikely to generate alternative frames independently (Hogarth, 1982; Kahneman & Tversky, 1984; Thaler, 1985). More open-ended procedures may clarify which frames people adopt spontaneously (Fischhoff, 1983; Frisch, 1993). Thus, teens concerned about relations with friends may not even think about how those choices affect subsequent decisions about smoking or sex.

Some of the most effective programs for reducing teen risk behaviors might be seen as attempts to eliminate such framing effects. These programs train teens in the cognitive skills needed to understand decisions and in the social skills needed to implement choices, by withstanding pressure to take unwise actions (Baron & Brown, 1991; Botvin, 1983; Dryfoos, 1990). Unfortunately, the empirical evaluations of these complex interventions reveal little about how they affect teens' decision-making processes en route to influencing their behaviors (Beyth-Marom *et al.*, 1991).

Structure

We used several coding schemes to ask whether the decisions described by teens had the structure assumed by decision theory: distinct options, clear consequences, and intervening uncertainties—as well as what forms these components take.

Options. The introductory segments of the interview defined decisions as choices among two or more options and saw that subjects could produce suitable examples. The decisions that subjects subsequently produced, however, typically focused on a single option. Indeed, the largest category of decisions (about 45%) involved statements of resolve, like "eat more healthfully" or "stop blaming others." In effect, they expressed the decision to stick with a previous choice, which might or might not have been implemented. These descriptions leave the alternative options unstated. Explicit alternatives were also absent from the next most common category (20% of descriptions), involving decisions about whether to do something (e.g., smoke cigarettes). Roughly equal numbers of decisions (about 15%) described two distinct options (e.g., whether to go to school or hang out with friends) or a set of options that seemed to have been identified, even if it was not described completely (e.g., which class to take, what to wear, with whom to have lunch). Very few decisions (5%) involved seeking or "designing" options (e.g., how to spend my free time, what to do about having fought with a friend). The three hard decisions that subjects described in detail were similarly focused on single options. There was, however, a much greater tendency to formulate them in terms of "whether to do X" (espe-

² One exception is the psychometric studies of risk (Fischhoff, Slovic, Lichtenstein, Read, & Combs, 1978; Slovic, 1987; Vlek & Stallen, 1980), looking for common dimensions of diverse risks. Although widely cited in debates over risk policy, these studies have had little impact on basic behavioral decision-making research.

TABLE 1
Option Structure

Decision	Structure					
	To do X	Whether to do	Two choices	Finite choices	What to do about	Missing ^a
Recent						
School	36.2%	29.0%	13.0%	13.0%	2.9%	5.8%
Freetime	31.9	18.8	17.3	20.3	4.3	7.2
Clothing	30.4	10.1	11.5	31.9	2.9	13.0
Peers	47.8	24.6	10.1	5.8	4.3	7.2
Health	55.0	20.2	1.4	2.9	2.9	17.4
Money	52.2	11.6	10.1	4.3	10.1	11.6
Parents	30.4	23.2	18.9	4.3	8.7	14.5
Hard						
First past	39.1	33.3	20.2	—	5.8	1.4
Second past	20.2	44.9	18.8	5.8	4.3	5.8
Current	2.9	44.9	23.2	13.0	13.0	2.9

^a Includes cases where respondents produced no answer or an uncodable one, or where the question was not asked due to a procedural error. These three cases constituted 27, 51, and 21% of missing responses, respectively.

Source: Fischhoff, Furby, Quadrel & Richardson (1991).

cially with the pending decisions, which were seldom described in terms of decisions “to do X”). The hard decisions were also more likely to involve two distinct options, rather than a set of possibilities. These results, and others, appear in Table 1.

These descriptions suggested some additional underlying mechanisms: (a) These teens rarely described their past decisions as having offered the possibility of creating new options, which might have gotten them out of difficult spots. (b) The chance to create options was more common with pending hard choices (suggesting that behavior looks more constrained in hindsight than in foresight). (c) Subjects’ hard decisions (like Choice B) were somewhat more likely to involve two clear options. (d) The more decisions that a subject described, the more likely those decisions were to be statements of resolve.

If these descriptions and our coding capture subjects’ everyday thinking, then their decision making focuses on the acceptance or rejection of a single option, or even just the reiteration of a previous choice. Paying undue attention to a focal option is seen in several phenomena demonstrated in laboratory experiments. These include neglect of opportunity costs, overemphasis of sunk costs, and various confirmation biases (e.g., Dawes, 1988; Fischhoff & Beyth-Marom, 1983; Klayman & Ha, 1987; Thaler, 1980, 1985). If people think the way they talk, then our subjects’ descriptions suggest that such phenomena are not only possible, but common.

Breadth. Decisions might be arranged along a continuum from *tactical*, one-time decisions regarding specific limited choices to *strategic* decisions, setting long-term, relatively irreversible policies. In order to see whether teens address policy issues spontaneously, we categorized their descriptions into (a) concrete choices, regarding specific situations; (b) limited policies, regarding a single repeated decision (e.g., ride my horse every day); (c) limited policies, regarding multiple decisions (e.g., save money for horse equipment); (d) general policies, with broad principles governing diverse circumstances (e.g., take better care of myself).

The vast majority of all decisions involved concrete, one-time choices. This was even more true for the detailed hard decisions than for the briefly described recent decisions (about 85% vs 75%). The clearest exception was decisions about health and money, 48 and 36% of which involved limited policies (e.g., what sort of diet or spending pattern to adopt). It is not hard to imagine general policy choices in the other areas (e.g., how to spend free time, what to do about homework, how to keep parents off one’s back). However, they were rarely produced. Although hard decisions provide an opportunity to reflect about the long term, it was not one that these teens seized.

In such situations, tactical situations add up to general policies. For example, each decision about doing homework (or hanging out with the crowd) partially defines a teen’s policies regarding schoolwork and responsibility (among other things). If one-time and re-

peated choices are made differently (Keren & Wagenaar, 1987), then results from single-choice experiments may be most pertinent to predicting real-world behavior, in situations that could be construed either way.

Consequences and uncertainties. Subjects chose one of their two past hard decisions for a fuller and more directed description. Among other things, they were asked to list six kinds of consequences. The first was “good things that you were certain would happen to you if you made that choice.” Subsequent requests dealt, in order, with uncertain good things, certain bad things, uncertain bad things, good things given up by making that choice, and bad things avoided by making that choice. The last two categories were called *indirect* consequences. Such open-ended, but directive questions show how subjects can think about these issues, an indirect indicator of how they actually do.

On average, subjects produced two good consequences of their chosen option, and one consequence in each other category. Three times as many subjects produced no certain bad consequence as produced no certain good one (27% vs 9%). Most subjects produced at least one good and one bad *uncertain* consequence (84%, 80%), indicating that they could think about uncertainty—if asked explicitly. Despite the direct query, quite a few subjects produced no indirect consequences, either good ones given up by the choice or bad ones avoided by it (29%, 30%). Here, too, there is confirmation of experimental results, showing the difficulty of thinking up opportunity costs, even when asked explicitly. This reduced sensitivity to indirect consequences applied equally to avoided costs and forgone benefits.

Process

We did not observe these decisions being made. Even for the pending hard decisions, we got only a mid-stream selection of issues. As a result, we also asked subjects to describe their decision-making processes. These descriptions are vulnerable to various biases (e.g., flawed introspection, self-serving biases). Even so, they may have a life of their own, if they reflect the stories that people tell themselves, about their decision making.

What made this a hard decision? Subjects produced two reasons, on average, for each of their three hard decisions. For the two past decisions, the most common concern was negative aspects of the chosen option, followed by negative features of the rejected option. They seldom mentioned positive features of either the chosen option or a rejected one (i.e., opportunity costs). Thus, they emphasized loss avoidance over benefit maximiza-

tion. Similarly, the difficulty of their current hard decisions was attributed much more often to both options having negative features than to both having positive features.

Subjects seldom mentioned uncertainty as a major source of difficulty, even though they had produced uncertain consequences in response to a previous question. Subjects never mentioned some potential reasons that we had included in our coding scheme because they are often mentioned as sources of decision-making difficulty: complexity (many options to choose from, many consequences to consider), time pressure, and personal inability to make decisions or make changes. In this light, either these are not the factors that make decisions difficult or people do not recognize their influence.

A subsequent question asked “what did you think about while making the decision?” It produced three issues on average. Similar patterns emerged (e.g., a focus on negative and concrete consequences). One difference was greater reference to uncertainty. However, it typically concerned how an event would be experienced (e.g., “what would it be like if I went to live with my mother” or “went to the other school”), rather than whether something would happen.

What did you do to help you think about your decision? For about one third of their past decisions, subjects reported doing nothing. Most of the remaining cases involved various ways of focusing their thinking or seeking information and support. Those who sought help went to friends and parents. A few subjects reported trying not to think. Even fewer reported reading something.

Subjects were much more likely to report active responses (beyond just thinking) for their current hard decisions, with a particular tendency to experimentation. Trying out different options (e.g., ways of dealing with parents) can reduce cognitive load (by acting out possibilities) and reduce uncertainty (by seeing how actions work and feel). Current experiments may be more visible than past ones because many plans do not actually get carried out. Thus, past decisions may provide better indicators of which experiments they get around to trying.

Do you wish that you had made a different choice? Only 15–20% of subjects said that they did. Their primary reason was not liking the option that they had selected, rather than wishing that they had taken the rejected option. Those who expressed no regret were also twice as likely to mention good aspects of what they had chosen, rather than bad aspects of what they had rejected. The lack of regret may reflect which hard

decisions subjects chose to describe, how well those decisions were made, or how they were reinterpreted in retrospect. These descriptions suggest a less balanced analysis of competing alternatives than that appearing in regret theories (Loomes & Sugden, 1983).

In response to other questions: (a) Subjects saw themselves as working about as hard as their peers at decision making. (b) Some 39% saw themselves as equally good decision makers. Those who made a distinction were twice as likely to describe themselves as better than average decision makers than as worse (42% vs 18%). (c) Individual differences in the difficulty of making decisions were attributed to the circumstances that teens faced and the social support that they enjoyed, rather than to personal characteristics. These beliefs contrast with the popular notion that teens view themselves as unique and invulnerable (e.g., Elkind, 1967; Quadrel *et al.*, 1993). They suggest resistance to the fundamental attribution error (Nisbett & Ross, 1980).

Summary

Taken at their word, these teens emerge as active, absorbed decision makers. After minimal instruction, they produced many examples of decisions in their lives. Their descriptions were sufficiently clear that we could code them from a variety of decision theoretic perspectives. The teens seemed eager to discuss their decisions and, as in Choice B, recognized their complexity. Nonetheless, their descriptions showed some potential problems. The focus on single options is reminiscent of several phenomena in the literature (e.g., neglect of opportunity costs), while the focus on negative consequences might be a corollary of loss aversion. The limited references to uncertainty about what will happen contrast with the focus on probability assessment in decision-making research. It might also provide weak general support for research showing problems with probability judgments, assuming that people do poorly at tasks that attract little of their attention. References to uncertainty centered on what (relatively certain) outcomes would be like. Understanding those uncertainties requires studying how people predict their tastes, how confident they are in those predictions, and how they resolve uncertain preferences (Fischhoff, 1991; Frisch, Jones & O'Brien, 1994; Kahneman, *in press*; Loewenstein, *in press*).

COMPONENTS OF DECISIONS

Letting people discuss whatever decisions are on their minds allows them to speak about relatively well-

worked topics. However, it produces such diverse decisions that only general, structural analyses are possible. It is difficult to study how people deal with the substance of particular domains. The following sections consider some possibilities for an open-ended focus on specific components of decisions.

Uncertainties

A natural question in considering teens' (or anyone else's) decisions about risks is how well they understand the size of those risks. Our literature review found many surveys of adolescents' risk perceptions. However, few had questions formulated precisely enough for the accuracy of the answers to be evaluated. As might be expected, most tasks used verbal quantifiers (e.g., likely, rarely), whose ambiguity is well known (Lichtenstein & Newman, 1967; Merz, Druzdzel, & Mazur, 1991; Wallsten, Budescu, Rapoport, Zwick, & Forsyth, 1986). Thus, these investigators had not yet begun to deal with the difficulties of eliciting quantitative estimates in unfamiliar units (Fischhoff & MacGregor, 1983; Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978; Poulton, 1989).

Less expected was the discovery of similar ambiguity in the descriptions of the risks being judged. Even had a precise response mode been used, the accuracy of those answers still could hardly be evaluated. Table 2 shows two questions from a large-sample survey, sponsored by the National Center for Health Statistics of the Centers for Disease Control. Following each question are alternative definitions of the quantity and intensity of that kind of risk behavior. If one believed that these exposures create any risk at all, then the magnitude of that risk would depend on what one inferred about these unspecified elements of the definition (among others). Subjects' risk judgments could be interpreted only if all subjects made the inferences—which investigators were then able to guess. The numbers in parentheses indicate the percentage of subjects who endorsed each alternative interpretation, after having answered the question on a previous page. These subjects were drawn from a relatively homogeneous population (juniors at an Ivy League college) The range of opinion shows some of the ambiguity in these questions and some of these students' linguistic norms and mental models.

The options in Table 2 were the investigators' own invention. Suggesting them might have artificially put some ideas in subjects' minds, while neglecting other ideas that were already there (thereby exaggerating or understating the degree of ambiguity). Quadrel (1990; Quadrel, Fischhoff, & Palmgren, 1994) adopted a more

TABLE 2
Interpretations of AIDS Risk Questions

How likely do you think it is that a person would get AIDS or the AIDS virus from sharing plates, forks or glasses with someone who has AIDS?

How did you interpret "sharing plates, forks, or glasses with someone who has AIDS?"

_____ Sharing utensils during a meal (e.g., passing them around, eating off one another's plates). (81.8%)

_____ Using the same utensils after they have been washed. (10.9%)

_____ I was uncertain about the interpretation. (5.8%)

_____ Multiple interpretations (1.5%)

Did you interpret "sharing plates, forks, or glasses with someone who has AIDS" as

_____ Occurring on a single occasion? (39.1%)

_____ Occurring on several occasions? (19.6%)

_____ Occurring routinely? (27.5%)

_____ I was uncertain about the interpretation. (12.3%)

_____ Multiple interpretations (1.4%)

How likely do you think it is that a person would get AIDS or the AIDS virus from having sex with a person who has AIDS?

How did you interpret "having sex with a person who has AIDS?"

_____ Having vaginal intercourse without a condom. (72.5%)

_____ Having vaginal intercourse with a condom. (4.3%)

_____ Having other kinds of sex. (6.5%)

_____ I was uncertain about the interpretation. (6.7%)

_____ Multiple interpretations. (8.0%)

Did you interpret "having sex with a person who has AIDS" as

_____ Occurring on a single occasion? (61.6%)

_____ Occurring on several occasions? (22.5%)

_____ Occurring on many occasions? (7.2%)

_____ I was uncertain about the interpretation. (5.1%)

_____ Multiple interpretations. (3.6%)

Note. Entries are the percentage of subjects (in a sample of 135 students at an Ivy League college) who reported having inferred each definition of the phrase when they had answered a question about the risk that it entailed. The response mode comprised five verbal quantifiers: definitely not possible, very unlikely, somewhat unlikely, somewhat likely, and very likely.

Source: Fischhoff (1994).

open-ended strategy: having teens think aloud as they answered risk-estimation questions that were deliberately ambiguous (at about the level of typical surveys). For example, "What is the probability that a person will have an accident while drinking and driving?" or "What is the probability that a person will get AIDS from having sex with someone who has AIDS?" Subjects were drawn from both low-risk settings (e.g., suburban high school teams and clubs) and high-risk ones (e.g., treatment homes).

Table 3 shows the coding framework for the risk factors that subjects produced, illustrated with examples from the drinking-and-driving question. Data analyses showed great variation in subjects' assumptions about the unstated elements of these risks (e.g., how much driving, what kind of driving, how serious an accident). As a result, these subjects were effectively answering different questions from one another and from any investigator's expectations. In these domains at least, it may be hard to learn very much about risk judgments without much sharper questions and answers. A follow-up experiment varied individual risk factors within

highly specified questions. Subjects there were most responsive to those factors that had been cited most frequently as relevant in the open-ended study.

Methodologically, these results suggest that well-specified questions are both needed and possible. Substantively, they suggest that teens have fairly robust intuitive theories regarding these risks, which emerge similarly in different tasks. The content of those beliefs suggest how they have processed past experiences with these risks and how they might make future decisions about them. For example, most subjects spontaneously discussed the frequency of most risk behaviors, indicating the salience of dose-response relationships. However, they did not ask about the frequency of sex as a determinant of the risks of pregnancy and of AIDS.

These particular results provide an additional perspective on a recurrent topic in experimental studies: how people estimate the cumulative probability of compound events (e.g., Bar-Hillel, 1974; Cohen, Chesnick & Haran, 1971; Keren, 1990). In structured tasks, we have observed particularly large underestimation of how quickly risks mount up through repeated sexual expo-

TABLE 3
Coding Framework

Framework element	Risk factor categories	Example variables drinking-and-driving question
Behavior	Dose Amount Potency Method	Amount of alcohol consumed
Other behaviors	Risk buffers Risk amplifiers Time-related Place-related	Amount of food eaten Other drugs consumed Night or day; day of the week Where alcohol was consumed
Actor	Physical Cognitive Social–psychological Material Spiritual Skill Character Age Gender Genetic history Status Luck Motivation Self other	Tolerance to alcohol Awareness of effects of alcohol Mood Wealth Faith Driving skill Responsible; mature
Context	Social General, cultural Family Peers, others	Drinking norms Family approval Peer approval
Risk outcome	Environmental Social reactions Personal effects Physical Psychological Cognitive–physiological Cognitive–psychological Material Accidents Lifestyle Complex Effects on others Behaviors Severity, type when measured	Road conditions Get in trouble Injury Worry, guilt Kill brain cells Can't think Lose car, lose license Get in a wreck while high Become a bum Get high Hurt your friends, family Use more, do heavier drugs

Source: Quadrel, Fischhoff & Palmgren (1996).

tures (Linville, Fischer, & Fischhoff, 1993; Shaklee & Fischhoff, 1990). The present, open-ended results suggest that the idea of cumulative risk does not even occur to many people in this domain (Luker, 1975). If an issue does not arise naturally, then people may do particularly poorly when asked to address it. There are many studies of intuitive beliefs in specific domains (e.g., Chi, Glaser & Farr, 1988; Furnham, 1988; Leventhal & Cameron, 1987; Morgan, 1993). Behavioral decision making might exploit them to identify and to understand the properties of the topics it happens to use for stimuli.

Consequences

A focal argument in youth policy is whether information “works” with adolescents. If education does not influence teen decision making, then there is a stronger case for implementing coercive policies. One approach to determining the role of information in teen decisions has been to predict those decisions as a linear function of the judged attractiveness of a preselected set of consequences, weighted by their judged likelihood. The moderate success of such models (e.g., Bauman, 1980;

Bauman, Fisher & Koch, 1988) could provide some weak support for the claim that teens' choices are under cognitive control. However, it has long been known, in behavioral decision-making research, that many weighting schemes will provide somewhat successful predictions, as long as the consequences span the space of relevant concerns—or are correlated to ones that do (Camerer, 1981; Dawes & Corrigan, 1974; Goldberg, 1968). The price for these practical successes is difficulty in distinguishing alternative specifications and, hence, in gaining much insight into cognitive processes (Hoffman, 1960).

In order to gain some additional purchase on which consequences matter to teens, we asked some teens to list the “things that might happen” if they accepted (or declined) opportunities to engage in six potentially risky behaviors (e.g., driving with a friend who has been drinking, skipping school to go to a mall) (Beyth-Marom, Austin, Fischhoff, Quadrel & Palmgren, 1993). We found very similar responses in matched groups of adults and adolescents, drawn from low-risk settings. In addition to sharing beliefs about these possibilities, these teens and adults were equally sensitive to how their task was formulated. Although accepting and declining a risky opportunity are logically complementary actions, they were not psychologically complementary. For example, engaging in a risk behavior generated more consequences overall than did avoiding it, but fewer indirect consequences (i.e., ones that do not occur as a result of the choice). Accepting a risky option produced four times as many bad consequences as good ones; however, rejecting it produced equal numbers of good and bad consequences. Subjects saw rejecting a risk as eroding their social standing much more than they saw accepting a risk as securing social approval. Another kind of sensitivity to question framing was that the prospect of repeatedly engaging in a risk behavior evoked many more bad consequences than did doing it just once.

These results suggest general decision-making processes. For example, one might infer that action and inaction evoke quite different perspectives, as do short-term and long-term time perspectives. Without explicit prompts, thinking about one perspective does not invoke the other. These results also elaborate or qualify conclusions reached in studies using structured methods. For example, they show another context within which opportunity costs are neglected, and that such neglect is less when rejecting risky options than when accepting them. Methodologically, these results suggest some possibilities for increasing the prediction ceiling on linear models (e.g., distinguishing between

accepting and rejecting an action, using teens' terminology for formulating consequences, adopting the appropriate time perspective).

In terms of the policy debates that motivated them, these studies show teens as having a higher capacity for cognitive performance than one might have expected. They also suggest some specifics on how to communicate best with teens. These two issues are interrelated. Poorly designed communications erode faith in teens, by making them responsible for our failures to speak clearly (we knew what we meant, why didn't they listen?). If teens are seen as cognitively competent, then there is more place for providing them with the sophisticated information needed for effective decision making, addressing the values that make them willfully choose behaviors that adults deplore, and alleviating circumstances that restrict their ability to act maturely.

Options

If only one thing is standardized in a structured study, it is likely to be the options. Life, however, does not always provide ordered sets of options. Indeed, some of the few studies of option generation have shown sensitivity to just how the tasks are posed by investigators (Gettys *et al.*, 1987). The differences between the perceived consequences of accepting and rejecting a risk behavior (Beyth-Marom *et al.*, 1993) suggest that an option need not even evoke its complement.

Some years ago, Lita Furby and I became interested in the confident, universal, and contradictory advice being offered to women, regarding how to reduce the risk of sexual assault (Morgan, 1986). Even if well intended, such advice can do women a serious disservice. Not only can it lead to ill-advised actions, but unfounded advice can increase the blame placed on those women whose chosen options do not succeed—because whatever they did was against the advice of some self-styled experts. Universal advice can misdirect women who do not accept its underlying trade-offs.

As part of a systematic look at this advice and the evidence supporting it (Fischhoff, 1992), we examined the scope, organization, and formulation of options among both experts and laypeople. In the advice literature, we found that options are often lumped into coarse categories with loaded labels, like active and passive.³ In order to understand lay perceptions, we

³ For example, “doing nothing” is a strategy, perhaps the appropriate one (e.g., for escaping an assailant who is primarily after the “satisfaction” of subduing a woman who has actively resisted him). Doing little about personal protection might be the right strategy for a woman who felt that her energies were best invested in social action.

posed a series of open-ended questions asking what a woman could do to protect herself in various circumstances (Furby, Fischhoff, & Morgan, 1990). Our subjects were 43–45 women drawn from each of three populations, students, middle-aged alumna of the same university, and working-class young mothers. Each woman produced 26 options on average. Although this is a large number of strategies for an individual to write on an open-ended questionnaire, it is still less than a tenth of the 300-plus different options produced by each group as a whole.⁴

We supplemented this list of options with ones produced by a group of men and a national sample of sexual assault experts, and additional ones found in a sample of 50 publications written for lay or professional audiences. In total, we found over 1100 different options. This is a bewildering number of possibilities for women or researchers to consider in any detail. In order to bring some order to this welter of options, we created a “strategy grammar.” In it, each strategy was described in the form, “Doing *action* X in order to achieve *intended effect* Y [which is believed to reduce the risk of rape].” The typology distinguishes further between the *stage* of an assault at which a strategy is directed (preventing an assault from occurring, preparing to react to an assault, or defending oneself during an assault) and the *level of action* involved (individual or societal).

In this view, a strategy may be ineffective either because it fails to produce the intended effect or because that effect does not deter sexual assaults. Making that distinction might help researchers to understand the effects of strategies (by looking separately at the two contingencies). Indeed, this categorization revealed a moderately consistent pattern in the otherwise confusing results from studies of the effectiveness of assault prevention strategies (Furby & Fischhoff, in press). Making that distinction might help women to generate options by prompting the search for ways to achieve these ends (Pitz, Sachs & Heerboth, 1980). Table 4 shows the most commonly mentioned strategies, for what women can do if assaulted (Stage III). All are individual level actions for situations where societal action has failed. The details in the table suggest the differing experiences of laymen, women, and experts. Responses of the three groups of women were remarkably similar and are combined here.

⁴ Options were coded as “different” if we could see how women might view them as having different probabilities of producing a set of focal consequences. Those consequences were the ones mentioned most frequently in another study, asking similar women to list important ones (Furby *et al.*, 1991).

The messiness of these options contrasts with the apparent simplicity of the few (perhaps two) well-defined options offered in many lab experiments. However, the simplicity of the options in experiments belies the complexity of the underlying option space. The options are often selected from an uncountable space of possibilities by a deliberately inscrutable procedure. (If subjects knew how stimuli were derived, then they might discern the purpose of the study and provide artifactually orderly responses). There may be value to studying how the nature of outcome spaces affects people’s thinking. For example, do people draw on different skills for arbitrarily created tasks and for messily specified ones? If so, then the generality of behavior observed with artificial stimuli might depend on their underlying outcome space (Evans, 1986; Hammond, 1966; Hogarth, 1981).

WHOLE DECISION STRATEGIES II: TRANSACTION ANALYSIS

Contingent Valuation

According to Executive Order 12291 (Bentkover, Covello, & Mumpower, 1985), cost-benefit analyses must be conducted for all significant federal actions. Where those actions affect the environment, that often means putting a price tag on goods not traded in any marketplace. These goods include such “intangible” effects as changes in genetic diversity, visibility at national parks, and discomfort (due to increased tropospheric ozone).

Fearing the neglect of these effects, some resource economists have resorted to asking respondents for dollar equivalents. They call their family of procedures “contingent valuation” mechanisms, requiring participants to act *as if* a market existed for the environmental change. These investigators have soldiered on for 15 or more years, despite membership in a profession with little faith in “expressed preferences” (Mitchell & Carson, 1989; Portney, 1994). Recently, they have achieved some notable institutional successes: a panel of (more and less) disinterested scholars, convened by NOAA (1994), has issued a cautious endorsement. Contingent valuation results influenced the monetary value of the Exxon Valdez settlement (to the extent that science played any role at all).⁵

A great attraction of contingent valuation is that it allows investigators to pose questions about anything that interests them. However, this flexibility exacts a

⁵ In the interests of disclosure, I was a consultant to the Department of Justice in this case.

TABLE 4
What People Think Women Are Likely to Do if Assaulted^a

Stage-level-intended effect ^b	Strategy	% Who mentioned strategy ^c (rank order in parentheses)		
		Females	Males	Experts
III-i-outside intervention	Scream or yell	66.7	56.8 (1)	69.8 (1/2)
III-i-establish distance or barrier	Run away	57.6	52.3 (2)	69.8 (1/2)
III-i-reduce assailant's propensity to rape	Talk way out of situation, reasoning	44.7	18.2 (8/10)	41.9 (4)
III-i-physically impede assailant	Fighting/physical self-defense techniques (e.g., martial arts or specific blows/hits)	43.9	38.6 (4)	53.5 (3)
III-i-physically impede assailant	Knee/kick him in the groin	42.4	47.7 (3)	11.6 (28/32)
II-i-reduce assailant's propensity to rape	Make yourself seem less attractive; do crude "unfeminine," gross, dirty things (e.g., vomit, urinate, defecate, pick nose, fake menstrual cramp, pass gas)	30.3	11.4 (17/18)	27.9 (8/9)
III-i-minimize assailant's propensity to rape	Cooperate with rapist/submit	28.8	15.9 (11/3)	34.9 (5)
III-i-physically impede assailant	Poke, jab, or claw rapist's eyes	26.5	27.3 (5)	20.9 (12/15)
III-i-physically impede assailant	Kick	25.8	18.2 (8/10)	25.6 (10)
III-i-reduce assailant's propensity to rape	State you have VD, are a virgin, menstruating, recovering from operation, diseased	22.7	13.6 (14/16)	32.6 (6)
III-i-physically impede assailant	Bite attacker, draw blood (e.g., bite his tongue or lips hard)	21.2	20.5 (6/7)	23.3 (11)
III-i-physically impede assailant	Use Mace, tear gas and chemical sprays	16.7	15.9 (11/13)	11.6 (28/32)

^a Strategies mentioned by at least 10% of respondents in each groups, in answer to questions R and S (females) or R (males and experts). See Furby, Fischhoff, and Morgan (1990) for verbatim questions. A difference of approximately 20% between females and either of the other groups is significant at $p < .01$.

^b Stage codes: I, prevent assault from occurring; II, prepare for reacting to an assault; III, defend yourself during a assault. Action level codes: s, societal action; i, individual action. Intended effect subscripts: p, perceived; a, actual. See Fischhoff, Furby, and Morgan (1990).

^c Tied ranks are indicated by giving the range of ranks that are tied (e.g., 8/10 means that there were three strategies all mentioned with the same frequency and thus tied for the 8th, 9th, and 10th slots).

Source: Furby, Fischhoff, & Morgan (1992).

price. Because they can be asked anything, respondents can take nothing for granted. As a result, investigators must first convey the complex details of a CV question and, then, allow respondents to articulate stable preferences. Many critics have worried about the cognitive demands imposed on CV respondents (e.g., Baron, 1994; Diamond & Hausman, 1994; Fischhoff, 1988, 1991; Kahneman & Knetsch, 1992; Schkade & Payne, 1994).

After reviewing the contingent valuation studies in a single domain (atmospheric visibility), Lita Furby and I settled on a rather different concern: that the intellectual demands of these tasks had overwhelmed the theoretical capacity of economics. We found enormous variability in task specification across studies, without clear guidelines for its resolution. Details emphasized by one investigator were often ignored entirely by another. Where one scientist feared misinterpretation,

another felt that the same detail went without saying. It seemed as though the analysis of existing markets, the customary task of economics, had not shown how to create new (contingent) ones.

In response, we created a framework, identifying the details required in a well-specified "transaction" (Fischhoff & Furby, 1988). It distinguished the *good* being offered, the *payment* being requested, and the *social context* within which the proposed transaction would be conducted. The good and payment have *substantive* properties, representing features that one might care about, and *formal* properties, defining the magnitude of those features (which should matter if one cares about them). For example, a substantive property of visibility is haze; its formal properties include the area and time over which a change would be experienced, as well as its probability of being achieved as promised. The social context includes what prece-

dents would be set, who legitimates the proposed transaction, and who else is being asked.

CV investigators typically have little interest in most of these details, caring only about a few focal features. However, there is no guarantee that omitted details will not be imputed by respondents, who need to assume something in order to make the task meaningful. Of course, respondents may be insensitive to these omissions and to much of what is included. Indeed, they may be content to give a general response to a general question. Such “gist” responses may be the best predictors of behavior in situations with low involvement. Unfortunately, CV studies require a much deeper level of processing: monetary evaluations for precisely described goods.

One obvious way to examine depth of processing is with manipulation checks, asking respondents how they interpreted their task. Recently, we quizzed people about the details of a brief evaluation task, asking about a river cleanup, which we had just read them over the phone (Fischhoff *et al.*, 1993). Many had forgotten (or never believed) essential details in our description, such as the extent of the cleanup, the likelihood of it actually being provided, and who else might pay. Their evaluations were more consistent when interpreted in terms of the task that they reported having answered, rather than the one that we had posed.

Insurance Decisions

Our constructive criticism has had some effect on CV practice (Pommerehne & Schwartz, 1995). It has also forced us to look again at the missing details in our own experimental tasks. Ones like Choice A require conveying many fewer details. On the other hand, such tasks leave subjects much greater latitude to read between the lines, because there are so many more details whose omission they might notice. This is, in fact, the same freedom of interpretation that we exploit when extrapolating lab results to the real world. Because we have said so little, we have great freedom in making claims about what in the world looks like our tasks.

Table 5 characterizes Choice A in terms of some elements in the framework that we developed for contingent valuation. It performs the same analysis for a second choice problem, identical to A, except that “certain loss” is replaced by “insurance premium.” Although the stakes are the same in the two problems, respondents are much more risk averse with the insurance formulation (Fischhoff, Slovic, & Lichtenstein, 1980; Hershey & Schoemaker, 1980).

We attributed this difference to one feature in the framework, the *payment context* (i.e., how the money

is extracted). This is a substantive issue, which might be expressed along the lines of “people don’t like certain losses, but do protect themselves when insurance is offered.” If true, this would mean that people use propositional rules, about how to behave in particular contexts, in addition to analytical rules, for aggregating vectors of dollars and probabilities.

Although brief, these choice descriptions do make several other features explicit. For example, both the good and the payment are in dollars; the choices are single-play, allowing no “iterations” (with opportunities to learn from experience or to have outcomes even out); the probabilities of the risk (.25) and the promised protection (1.0) are given. A single question mark indicates a missing feature that subjects might have imputed, but probably did so similarly for both choices. For example, the hypotheticality might have influenced subjects’ interpretations of these probabilities. Our interpretation of the different preferences for the two choices need not change if subjects made similar guesses about these features, even if we do not know what those guesses are.

Life is more complicated if subjects made different assumptions in the two conditions. In that case, we do not know what questions subjects answered nor why they answered differently. Double question marks indicate such cases. For example, the source of the threatened \$200 loss may seem different when insurance is available than when a certain loss will avert it. If the source of a loss matters, then inferences about this missing detail may influence choices. While the context (premium/sure loss) may have been the cue to these inferences, it would be an incomplete explanation of subjects’ differential desire for these two kinds of protection.

The same could be said for possibly different interpretations of the unstated social context features: who provides the coverage (and is party to the contract), what constraints are there on possible offers (and opportunities to take advantage of subjects), what precedents are being set (relative to subjects’ self-concept or their relationship with the coverage provider), and what social norms might legitimate (or proscribe) such a deal. If they affect choices, these interpretations would elaborate the psychology of “certain loss” and “premium.” They would provide a basis for linking this preference reversal to the complex settings within which decisions about protective behavior are actually made.

A framework is not a theory. However, it can identify issues that may warrant theorizing and might be studied across contexts. Thus, one might examine people’s

TABLE 5
Transaction Analysis Insurance/Preference

	Preference ^a	Insurance ^b
Good		
Attribute	\$	\$
Source of change	??	??
Extent	one play?	one play?
Certainty of provision (probability)		
Receiving needed compensation	1.00?	1.00?
Needing compensation	.25?	.25?
Payment		
Attribute	\$	\$
Context	premium	sure loss
Constituency	??	??
Certainty of provision (probability)	1.00	1.00
Social context		
Providers	??	??
Others involved	Exp. group, family, friends?	Exp. group, family, friends?
Determined by	self?	self?
Iterations	none	none
Constraints	??	??
Precedents	??	??
Legitimacy	??	??

Note. ?, uncertain value; ??, uncertain, possibly different value in two conditions.

^a Choice A with "certain loss" replaced by "insurance premium."

^b Choice A at the beginning of the article.

Source: Fischhoff (1993).

sensitivity to the probability of actually receiving a promised good, to which other people observe their choice, or to the schedule for making payments. Insofar as these issues have been studied in their own right, one could then exploit the relevant chapters in psychology, economics, or sociology. Doing so might inform a general theory of context and limit speculations about the boundary conditions on individual studies.

CONCLUSION

Is This Science Good for the Real World?

Although designed with real-world problems in mind, these studies are still quite a ways from improving people's lives. Creating workable interventions requires detailed design work, followed by rigorous empirical evaluation. Baron and Brown (1991) show the work needed for one form of intervention, training in decision making (see also Beyth-Marom *et al.*, 1985). Once proven, an intervention must be sold to those controlling access to applied settings (e.g., school administrators, federal regulators, police officials).

These activities require a focus, a work rhythm, and a skill set quite different than those for managing basic research programs. Rather than retooling themselves,

basic researchers might adopt the more modest objective of making their results available to those better situated to acting on them. Unfortunately, providing that access requires considerable repetition. The same message needs to be delivered in person and in writing to a variety of practitioner audiences, tailored to the circumstances and language of each. They may even want to meet and query the individual behind the words. Considerable tending is needed to ensure that the message is heard consistently, in a world where it is convenient to believe one thing or another about human behavior (Fischhoff, 1990).

A more modest interim aspiration is to clarify bits of the long path between the abstractions of the lab and the messiness of the world. Doing so might help basic researchers take responsible steps toward practicality, without promoting lab results too quickly and confidently. It might help practitioners refine existing programs whose cognitive components are underdeveloped. As mentioned, adults often despair of teens (just as technical experts often despair of laypeople) for doing the wrong things "despite being told the facts." Yet, the telling is often conducted with little attention to what people need to know, what they know already, how they interpret messages, and how they process

information—not to mention how they might legitimately extract different conclusions from the same set of facts (Fischhoff, Bostrom, & Quadrel, 1993; Merz *et al.*, 1993). The research record shows how we have discovered these issues and partial solutions, in our communications with subjects.

If I were to speculate about the greatest real-world impact of this research, I would place it in this area. Beliefs about competence are important insofar as more competent people are afforded greater autonomy. Decision-making research has sometimes been funded by authorities hoping that psychologists could “do something” about a problematic group of people (e.g., citizens who dislike a technology, adolescents who take health risks). Two early, and oft-repeated, research messages have been: (a) it takes systematic evidence (and not just anecdotal observation) to make responsible statements about others’ behavior; and (b) both experts and laypeople show complex patterns of strengths and weaknesses (Fischhoff, 1977, 1990; Fischhoff, Slovic, & Lichtenstein, 1983; Quadrel *et al.*, 1993). I think that these efforts have contributed to a more mutually respectful atmosphere, in arenas where there are strong incentives to bash the public (and to distrust the experts) (National Research Council, 1989).

It is, of course, hard to evaluate a claim of having provided insight to a wide audience. Even invitations to speak and citations in the national press are but circumstantial evidence. Psychology may be no more than a convenient place holder, showing that the public has been represented, even if the details of its concerns are ignored. A somewhat more direct, but still ambiguous contribution is to provide components of larger programs. For example, we developed a set of open-ended procedures to describe and improve people’s “mental models” of environmental hazards (Bostrom, Fischhoff, & Morgan, 1992; Morgan *et al.*, 1992). It has influenced the design of two science museum exhibits on climate change and the re-design of EPA’s *Citizen’s Guide to Radon*. We used it to design a brochure on the health effects of electromagnetic fields, over 100,000 copies of which we have sold (at cost) (Morgan, 1989). A large chemical company has incorporated elements of this approach in its communications, extending it to topics like corporate restructuring. However, even though we designed our procedures for practicality, our colleagues in the company had to improvise extensively in order to use them. What they could use fairly directly was our articulation of a philosophy promoting an open attitude toward their audiences (Fischhoff, 1995). In each case, though, our research was but one piece of a larger process, whose overall impact is poorly documented.

Is the Real World Good for This Science?

It is difficult to study simultaneously how people perceive the world and what they do with those perceptions. A standard strategy in cognitive psychology is to standardize the stimuli, so as to gain access to the process. In this light, subjects become something like battery-raised hens, placed in very similar environments, in hopes of achieving very similar outcomes. The complex intervening processes are clarified whenever changes in individual input factors (e.g., temperature, crowding, diet) produce predictable changes in outputs.

With poultry, this strategy leads to harsh practices. For example, the destabilization of normal social relations may prompt trimming chicken’s beaks or putting on little red spectacles, in order to keep them from pecking one another. Although psychologists are prevented (by ethics and regulations) from abusing subjects with strange procedures, we are not protected from deluding ourselves. Live with a task long enough and it may become more real than the setting from which it was abstracted.

Tasks like Choice A run the risk of what might be called the “curse of context”: We would like to interpret responses as reflecting deep-seated values, of the sort that come from a lifetime of intense involvement with real-world decisions. Yet, we set minimalist problems before subjects and expect them to resist the temptation to impute a context. We may also be victims of a “curse of cleverness”: We pride ourselves on devising just the right tasks for evaluating competing theories of human behavior, tasks that our predecessors were unable to concoct. Then, we expect subjects immediately to discern the structure of these tasks and decide what response is right for them.

The studies reported here “turn subjects loose” within (something like) the domains that more structured studies hope to illuminate. If subjects behave similarly when less constrained, there is added reason for confidence in existing results. If subjects behave differently or interpret tasks in unexpected ways, additional experimental controls may be needed in order to obtain the desired standardization. Although troubling in the short run, such problems may open opportunities for new structured tasks, studying the unexpected behaviors systematically. Assuming that these phenomena can be domesticated as part of experimental science, open-ended approaches can be one element in the anarchic process by which its intellectual capital is created (Feyerabend, 1975). Subjects in these tasks might be compared to free-range chickens, relative to the battery-raised hens of experimental tasks. None-

theless, even these studies are a long way from real-life decisions. They provide, for example, little in the way of social and emotional pressure. The free-range chicken is still a poor excuse for a bird.

Choosing research strategies is a gamble with a poorly structured option space. Narrowing the focus of tasks reduces one's ability to identify surprising phenomena, while increasing one's ability to make science out of what is found. Pressing harder for applied impact increases the chances of getting practitioners' attention, while decreasing the chance of having something to say—if it reduces the freedom of inquiry needed for basic applied research. Pressing too feebly for applications risks forfeiting the intellectual challenges thrown off by real-world problems. Although practitioners may misdiagnose their situation, there is often something to be understood when they perceive a problem. Doing a little applied basic research along the way may be a way of returning the favor.

REFERENCES

- Arnett, J. (1993). Reckless behavior in adolescence: A developmental perspective. *Developmental Review*, 12, 339–373.
- Baddeley, A. (1979). Applied cognitive and cognitive applied research. In L. G. Nilsson (Ed.), *Perspectives on memory research*. Hillsdale, NJ: Erlbaum.
- Bar-Hillel, M. (1974). Similarity and probability. *Organizational Behavior and Human Performance*, 11, 277–282.
- Baron, J. (1994). *Rationality and invariance: Response to Schuman*. Unpublished manuscript, presented at DOE/EPA Workshop on Using Contingent Valuation to Measure Non-Market Values.
- Baron, J., & Brown, R. (Eds.). (1991). *Teaching decision making to adolescents*. Hillsdale, NJ: Erlbaum.
- Bauman, K. E. (1980). *Predicting adolescent drug use: Utility structure and marijuana*. New York: Praeger.
- Bauman, K. E., Fisher, L. A., & Koch, G. G. (1988). External variables, subjective expected utility and adolescent behavior with alcohol and cigarettes. *Journal of Applied Social Psychology*, 19, 789–804.
- Bentkover, J., Covello, V. T., & Mumpower, J. (Eds.) (1985). *Benefits assessment: The state of the art*. Amsterdam: Reidel.
- Beyth-Marom, R., Austin, L., Fischhoff, B., Palmgren, C., & Quadrel, M. J. (1993). Perceived consequences of risky behaviors. *Developmental Psychology*, 29, 549–563.
- Beyth-Marom, R., Fischhoff, B., Quadrel, M. J., & Furby, L. (1991). Teaching adolescents decision making. In J. Baron & R. Brown (Eds.), *Teaching decision making to adolescents* (pp. 19–60). Hillsdale, NJ: Erlbaum.
- Beyth-Marom, R., Dekel, S., R. Gombo, R., & M. Shaked, M. (1985). *An elementary approach to thinking under uncertainty*. Hillsdale, NJ: Erlbaum.
- Bostrom, A., Fischhoff, B., & Morgan, M. G. (1992). Characterizing mental models of hazardous processes: A methodology and an application to radon. *Journal of Social Issues*, 48(4), 85–100.
- Botvin, G. J. (1983). *Life skills training*. New York: Smithfield Press.
- Camerer, C. (1981). General conditions for the success of bootstrapping models. *Organizational Behavior and Human Performance*, 27, 411–422.
- Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Cohen, J., Chesnick, E. I., & Haran, D. (1971). Evaluation of compound probabilities in sequential choice. *Nature*, 32, 414–416.
- Dawes, R. M. (1988). *Rational choice in an uncertain world*. San Diego, CA: Harcourt Brace Jovanovich.
- Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81, 95–106.
- Diamond, P. A., & Hausman, J. R. (1994). Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives*, 8(4), 45–64.
- Dryfoos, J. G. (1990). *Adolescents at risk: Prevalence and prevention*. New York: Oxford Univ. Press.
- Elkind, D. (1967). Egocentrism in adolescence. *Child Development*, 38, 1025–1034.
- Ericsson, A., & Simon, H. A. (1984). *Verbal reports as data*. Cambridge, MA: MIT Press.
- Evans, J. St. B. T. (1986). *Deduction*. Hillsdale, NJ: Erlbaum.
- Feyerabend, P. (1975). *Against method*. New York: NLB (Schocken).
- Fischhoff, B. (1995). Risk perception and communication unplugged: Twenty years of process. *Risk Analysis*, 15(2), 137–145.
- Fischhoff, B. (1994). What forecasts (seem to) mean. *International Journal of Forecasting*, 10, 387–403.
- Fischhoff, B. (1993). Transaction analysis: A framework and an application to insurance decisions. *Journal of Risk and Uncertainty*, 7, 53–69.
- Fischhoff, B. (1992). Giving advice: Decision theory perspectives on sexual assault. *American Psychologist*, 47, 577–588.
- Fischhoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist*, 46, 835–847.
- Fischhoff, B. (1990). Psychology and public policy: Tool or tool maker? *American Psychologist*, 45, 57–63.
- Fischhoff, B. (1988). Specifying value measurements. In B. Driver, G. Peterson, and R. Gregory (Eds.), *Evaluating amenity resources* (pp. 107–116). New York: Venture.
- Fischhoff, B. (1983). Predicting frames. *Journal of Experimental Psychology: Learning Memory & Cognition*, 9, 113–116.
- Fischhoff, B. (1977). Cost-benefit analysis and the art of motorcycle maintenance. *Policy Sciences*, 8, 177–202.
- Fischhoff, B. & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, 90, 239–260.
- Fischhoff, B., Bostrom, A., & Quadrel, M. J. (1993). Risk perception and communication. *Annual Review of Public Health*, 14, 183–203.
- Fischhoff, B., & Furby, L. (1988). Measuring values: A conceptual framework for interpreting transactions. *Journal of Risk and Uncertainty*, 1, 147–184.
- Fischhoff, B., Furby, L., Quadrel, M. J., & Richardson, E. (1991). Adolescents' construal of choice situations: Are their decisions our "decisions"? Unpublished manuscript.
- Fischhoff, B., & MacGregor, D. (1983). Judged lethality: How much people seem to know depends upon how they are asked. *Risk Analysis*, 3, 229–236.
- Fischhoff, B., Quadrel, M. J., Kamlet, M., Loewenstein, G., Dawes, R., Fischbeck, P., Klepper, S., Leland, J., & Stroh, P. (1993). Em-

- bedding effects: Stimulus representation and response modes. *Journal of Risk and Uncertainty*, 6, 211–234.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1980). Knowing what you want: Measuring labile values. In T. Wallsten (Ed.), *Cognitive processes in choice and decision behavior* (pp. 117–141). Hillsdale, NJ: Erlbaum.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1983). The “public” vs. the “experts”: Perceived vs. actual disagreement about the risks of nuclear power. In V. Coviello, G. Flammm, J. Rodericks, & R. Tardiff (Eds.), *Analysis of actual vs. perceived risks* (pp. 235–249). New York: Plenum.
- Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, 8, 127–152.
- Frisch, D. (1993). Reasons for framing. *Organizational Behavior and Human Decision Processes*, 54, 399–429.
- Frisch, D., Jones, K. T., & O'Brien, M. (1994). *Sources of difficulty in real-life decisions*. Unpublished manuscript, University of Oregon.
- Furby, L., & Fischhoff, B. (in press). Rape self-defense strategies: A review of their effectiveness. *Victimology*.
- Furby, L., Fischhoff, B., & Morgan, M. (1992). Preventing rape: How people perceive the options of defending oneself during an assault. In E. C. Viano (Ed.), *Critical issues in victimology: International perspectives* (pp. 174–189). New York: Springer Publishers.
- Furby, L., Fischhoff, B., & Morgan, M. (1990). Preventing rape: How people perceive the options for assault prevention. In E. C. Viano (Ed.), *The victimology research handbook* (pp 227–259). New York: Garland Publishing, Inc.
- Furby, L., Fischhoff, B., & Morgan, M. (1991). Rape prevention and self-defense: At what price? *Women's Studies International Forum*, 14(1/2), 49–62.
- Furnham, A. F. (1988). *Lay theories*. London: Pergamon Press.
- Gallie, W. B. (1964). *Philosophy and the historical understanding*. London: Chatto & Windus.
- Gergen, K. J. (1994). *Realities and relationships*. Cambridge, MA: Harvard Univ. Press.
- Gettys, C. F., Pliske, R. M., Manning, C., & Casey, J. T. (1987). An evaluation of human act-generation performance. *Organizational Behavior and Human Decision Processes*, 39, 23–51.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard Univ. Press.
- Goldberg, L. R. (1968). Simple models or simple processes? *American Psychologist*, 23, 483–496.
- Hammond, K. R. (1966). *The psychology of Egon Brunswick*. New York: Holt, Rinehart & Winston.
- Hershey, J. R., & Schoemaker, P. J. H. (1980). Risk taking and problem context in the domain of losses: An expected utility analysis. *Journal of Risk and Insurance*, 47, 111–132.
- Hexter, J. H. (1971). *The history primer*. New York: Basic Books.
- Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin*, 57, 116–131.
- Hogarth, R. M. (Ed.) (1982). *New directions for methodology of social and behavioral science: Question framing and response consistency*. San Francisco: Jossey-Bass.
- Hogarth, R. M. (1981). Beyond discrete biases. *Psychological Bulletin*, 90, 197–217.
- Kahneman, D. & Knetsch, J. (1992). Valuing public goods: The pursuit of moral satisfaction. *Journal of Environmental Economics & Management*, 22, 57–70.
- Kahneman, D. (in press). This volume.
- Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341–350.
- Keren, G., & Wagenaar, W. A. (1987). Violations of utility theory in unique and repeated gambles. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 13, 387–391.
- Klayman, J., & Ha, Y-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211–228.
- Kren, G. H., & Rapoport, L. H. (Eds.) (1976). *Varieties of psychohistory*. New York: Springer.
- Kuhn, T. S. (1962). *The structure of scientific revolution*. Chicago: University of Chicago Press.
- Lakatos, I. (1970). Falsification and scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of scientific knowledge*. Cambridge: Cambridge Univ. Press.
- Leventhal, H., & Cameron, L. (1987). Behavioral theories and the problem of compliance. *Patient Education and Counseling*, 10, 117–138.
- Lichtenstein, S., & Newman, J. R. (1967). Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science*, 9, 563–564.
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 551–578.
- Linville, P. W., Fischer, G. W., & Fischhoff, B. (1993). AIDS risk perceptions and decision biases. In J. B. Pryor & G. D. Reeder (Eds.), *The social psychology of HIV infection* (pp. 5–38). Hillsdale, NJ: Erlbaum.
- Loewenstein, G. (in press). This volume.
- Loomes, G., & Sugden, R. (1983). Regret theory: An alternative approach to rational choice under uncertainty. *Economic Journal*, 92, 805–824.
- Luker, K. (1975). *Taking chances: Abortion and the decision not to contracept*. Berkeley: University of California Press.
- Merz, J., & Druzzdel, M. & Mazur, D. J. (1991). Verbal expressions of probability in informed consent litigation. *Medical Decision Making*, 11, 273–281.
- Merz, J. F., Fischhoff, B., Mazur, D. J., & Fischbeck, P. S. (1993). A decision-analytic approach to developing standards of disclosure for medical informed consent. *Journal of Products and Toxics Liabilities*, 15, 191–215.
- Mitchell, R. C., & Carson, R. T. (1989). *Using surveys to value public goods: The contingent valuation method*. Washington, DC: Resources for the Future.
- Morgan, M. G. (1993). Risk analysis and management. *Scientific American*, 269, 32–41.
- Morgan, M. G. (1989). *Electric and magnetic fields from 60 Hz electric power: What do we know about possible health risks*. Department of Engineering and Public Policy, Carnegie Mellon University.
- Morgan, M. (1986). Conflict and confusion: What rape prevention experts are telling women. *Sexual Coercion and Assault*, 1(5), 160–168.
- Morgan, M. G., Fischhoff, B., Bostrom, A., Lave, L., & Atman, C. J. (1992). Communicating risk to the public. *Environmental Science and Technology*, 26, 2048–2056.

- National Oceanographic & Atmospheric Administration (1994). *Federal Register*, 59(5), 1062–1091.
- National Research Council. (1989). *Risk perception and communication*. Washington, DC: The Council.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice–Hall.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Pitz, G. F., Sachs, N. J., & Heerboth, J. (1980). Procedures of eliciting choices in the analysis of individual decision. *Organizational Behavior and Human Performance*, 26, 396–408.
- Pommerehne, W. & Schwarz, N. (Eds.). (1995). *Determining the value of nonmarketed goods*. New York: Plenum.
- Portney, P. R. (1994). The contingent valuation debate. *Journal of Economic Perspectives*, 8(4), 3–17.
- Poulton, E. C. (1989). *Bias in quantifying judgment*. Hillsdale, NJ: Erlbaum.
- Quadrel, M. J. (1990). *Elicitation of adolescents' risk perceptions: Qualitative and quantitative dimensions*. Ph.D. dissertation, Carnegie Mellon University.
- Quadrel, M. J., Fischhoff, B., & Davis, W. (1993). Adolescent (in)vulnerability. *American Psychologist*, 48, 102–116.
- Quadrel, M. J., Fischhoff, B., Fischhoff, M., & Halpern S. (1994). *Calibration of adolescents' and adults' confidence judgments*. Submitted for publication.
- Quadrel, M. J., Fischhoff, B., & Palmgren, C. (1996). *Adolescents' definitions of risk behaviors*. Submitted for publication.
- Schkade, D., & Payne, J. W. (1994). How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics and Management*, 26, 88–109.
- Shaklee, H., & Fischhoff, B. (1990). The psychology of contraceptive surprises: Judging the cumulative risk of contraceptive failure. *Journal of Applied Psychology*, 20, 385–403.
- Slovic, P. (1987). Perception of risk. *Science*, 236, 280–285.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1, 39–60.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4, 199–214.
- Vlek, C. A. J., & Stallen, P. J. (1980). Rational and personal aspects of risk. *Acta Psychologica*, 45, 273–300.
- Wallsten, T., Budescu, D. V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General*, 115, 348–365.

Received: October 4, 1995