IMPROVING RISK COMMUNICATION

Committee on Risk Perception and Communication

Commission on Behavioral and Social Sciences and Education
Commission on Physical Sciences, Mathematics, and Resources

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1989

Appendix C
Risk: A Guide to Controversy

BARUCH FISCHHOFF

FOREWORD BY THE COMMITTEE

This appendix was written by Baruch Fischhoff to assist in the deliberations of the National Research Council's Committee on Risk Perception and Communication. It describes in some detail the complications involved in controversies over managing risks in which risk perception and risk communication play significant roles. It addresses these issues from the perspective of many years of research in psychology and other disciplines. The text of the committee's report addresses many of the same issues, and, not surprisingly, many of the same themes, although the focus of the report is more general. The committee did not debate all points made in the guide. Even though this appendix represents the views of only one member, the committee decided to include it because we believe the guide to be a valuable introduction to an extremely complicated literature.

PREFACE

This guide is intended to be used as a practical aid in applying general principles to understanding specific risk management controversies and their associated communications. It might be thought of as a user's guide to risk. Its form is that of a "diagnostic guide," showing participants and observers how to characterize risk controversies along five essential dimensions, such as "What are the (psychological) obstacles to laypeople's understanding of risks?" and "What are the limits to scientific estimates of riskiness?" Its style is intended to be nontechnical, thereby making the scientific literature on risk accessible to a general audience. It is hoped that the guide will help make risk controversies more comprehensible and help citizens and professional risk managers play more effective roles in them.

The guide was written for the committee by one of its members. Its substantive contents were considered by the committee in the course of its work, either in the form of published articles and books circulated to other committee members or in the form of issues deliberated at its meetings. As a document, the guide complements the conclusions of the committee's report.

CONTENTS

I
INTRODUCTION

Risk management is a complex business. So are the controversies that it spawns. And so are the roles that risk communication must perform. In the face of such complexity, it is tempting to look for simplifying assumptions. Made explicit, these assumptions might be expressed as broad statements of the form, "what people really want is . . ."; "all that laypeople can understand is . . ."; or "industry's communicators fail whenever they. . . ." Like other simplifications in life, such assumptions provide some short-term relief at the price of creating long-term complications. Overlooking complexities eventually leads to inexplicable events and ineffective actions.

On one level this guide might be used like a baseball scorecard detailing the players' identities and performance statistics (perhaps along with any unique features of the stadium, season, and rivalry). Like a ballgame, a risk controversy should be less confusing to spectators who know something about the players and their likely behavior under various circumstances. Thus, experts might respect the public more if they were better able to predict its behavior, even if they would prefer that the public behave otherwise. Similarly, understanding the basics of risk analysis might make disputes among technical experts seem less capricious to the lay public.

More ambitiously, such a guide might be used to facilitate effective action by the parties in risk controversies, like the *Baseball Abstract* (James, 1988) in the hands of a skilled manager. For example, the guide discusses how to determine what the public needs to know in particular risky situations. Being able to identify those needs may allow better focused risk communication, thereby using the public's limited time wisely and letting it know that the communicators really care about the problems that the public faces. Similarly, understanding the ethical values embedded in the definitions of ostensibly technical terms (e.g., risk, benefit, voluntary) can allow members of the public to ask more penetrating questions about whose interests a risk analysis serves. Realizing that different actors use a term like "risk" differently should allow communicators to remove that barrier to mutual understanding.

USAGE

The guide's audience includes all participants and observers of risk management episodes involving communications. Its intent is to help government officials preparing to address citizens' groups, industry representatives hoping to site a hazardous facility without undue controversy, local activists trying to decide what information they need and whether existing communications meet those needs, and academics wondering how central their expertise is to a particular episode.

The premise of the guide is that risk communication cannot be understood in isolation. Rather, it is one component of complex social processes involving complex individuals. As a result, this fuller context needs to be understood before risk communication can be effectively transmitted or received. That context includes the following elements and questions:

- *The Science.* What is the scientific basis of the controversy? What kinds of risks and benefits are at stake? How well are they understood? How controversial is the underlying science? Where does judgment enter the risk estimation process? How well is it to be trusted?
- *Science and Policy.* In what ways does the nature of the science preempt the policymaking process (e.g., in the definition of key terms, like "risk" and "benefit"; in the norms of designing and reporting studies)? To what extent can issues of fact and of value be separated?
- *The Nature of the Controversy.* Why is there a perceived need for risk communication? Does the controversy reflect just a disagreement about the magnitude of risks? Is controversy over risk a surrogate for controversy over other issues?
- *Strategies for Risk Communication.* What are the goals of risk communication? How can communications be evaluated? What burden of responsibility do communicators bear for evaluating their communications, both before and after dissemination? What are the alternatives for designing risk communication programs? What are the strengths and weaknesses of different approaches? How can complementary approaches be combined? What nonscientific information is essential (e.g., the mandates of regulatory agencies, the reward schemes of scientists)?
- *Psychological Principles in Communication Design.* What are the behavioral obstacles to effective risk communication? What kinds

of scientific results do laypeople have difficulty understanding? How does emotion affect their interpretation of reported results? What presentations exacerbate (and ameliorate) these problems? How does personal experience with risks affect people's understanding?

SOME CAUTIONS

A diagnostic guide attempts to help users characterize a situation. To do so, it must define a range of possible situations, only one of which can be experienced at a particular time. As a result, the attempt to make one guide fit a large universe of risk management situations means that readers will initially have to read about many potential situations in order to locate the real situation that interests them. With practice, users should gain fluency with a diagnostic approach, making it easier to characterize specific situations. It is hoped that the full guide will be interesting enough to make the full picture seem worth knowing.

At no time, however, will diagnosis be simple or human behavior be completely predictable. All that this, or any other, diagnostic guide can hope to do is ensure that significant elements of a social-political-psychological process are not overlooked. For a more detailed treatment, one must look to the underlying research literature for methods and results. To that end, the guide provides numerous references to that literature, as well as some discussion of its strengths and limitations.

To the extent that a guide is useful for designing and interpreting a communication process, it may also be useful for manipulating that process. In this regard, the material it presents is no different than any other scientific knowledge. This possibility imposes a responsibility to make research equally available to all parties. Therefore, even though this guide may suggest ways to bias the process, it should also make it easier to detect and defuse such attempts.

II
THE SCIENCE

By definition, all risk controversies concern the risks associated with some hazard. However, as argued in the text of the report and in this diagnostic guide, few controversies are only about the size of those risks. Indeed, in many cases, the risks prove to be a side issue, upon which are hung disagreements about the size and distribution of benefits or about the allocation of political power in a society. In all cases, though, some understanding of the science of risk is needed, if only to establish that a rough understanding of the magnitude of the risk is all that one needs for effective participation in the risk debate. Following the text, the term "hazard" is used to describe any activity or technology that produces a risk. This usage should not obscure the fact that hazards often produce benefits as well as risks.

Understanding the science associated with a hazard requires a series of essential steps. The first is identifying the scope of the problem under consideration, in the sense of identifying the set of factors that determine the magnitude of the risks and benefits produced by an activity or technology. The second step is identifying the set of widely accepted scientific "facts" that can be applied to the problem; even when laypeople cannot understand the science underlying these facts, they may at least be able to ensure that such accepted wisdom is not contradicted or ignored in the debate over a risk. The third step in understanding the science of risk is knowing how it depends on the educated intuitions of scientists, rather than on accepted hard facts; although these may be the judgments of trained experts, they still need to be recognized as matters of conjecture that are both more likely to be overturned than published (and replicated) results and more vulnerable to the vagaries of psychological processes.

WHAT ARE THE BOUNDS OF THE PROBLEM?

The science learned in school offers relatively tidy problems. The typical exercise in, say, physics gives all the facts needed for its solution and nothing but those facts. The difficulty of such problems for students comes in assembling those facts in a way that provides the right answer. (In more advanced classes, one may have to bring some general facts to bear as well.)

The same assembly problem arises when analyzing the risks and benefits of a hazard. Scientists must discover how its pieces fit together. They must also figure out what the pieces are. For example, what factors can influence the reliability of a nuclear power plant? Or, whose interests must be considered when assessing the benefits of its operation? Or, which alternative ways of generating electricity are realistic possibilities?

The scientists responsible for any piece of a risk problem must face a set of such issues before beginning their work. Laypeople trying to follow a risk debate must understand how various groups of scientists have defined their pieces of the problem. And, as mentioned in the report, even the most accomplished of scientists are laypeople when it comes to any aspects of a risk debate outside the range of their trained expertise.

The difficulties of determining the scope of a risk debate emerge quite clearly when one considers the situation of a reporter assigned to cover a risk story. The difficult part of getting most environmental stories is that no one person has the entire story to give. Such stories typically involve diverse kinds of expertise so that a thorough journalist might have to interview specialists in toxicology, epidemiology, economics, groundwater movement, meteorology, and emergency evacuation, not to mention a variety of local, state, and federal officials concerned with public health, civil defense, education, and transportation.

Even if a reporter consults with all the relevant experts, there is no assurance of complete coverage. For some aspects of some hazards, no one may be responsible.

For example, no evacuation plans may exist for residential areas that are packed "hopelessly" close to an industrial facility. No one may be capable of resolving the jurisdictional conflicts when a train with military cargo derails near a reservoir just outside a major population center. There may be no scientific expertise anywhere for measuring the long-term neurological risks of a new chemical.

Even when there is a central address for questions, those occupying it may not be empowered to take firm action (e.g., banning or exonerating a chemical) or to provide clear-cut answers to personal questions (e.g., "What should I do?" or "What should I tell my children?"). Often those who have the relevant information refuse to divulge it because it might reveal proprietary secrets or turn public opinion against their cause.

Having to piece together a story from multiple sources, even recalcitrant ones, is hardly new to journalists. What is new about many environmental stories is that no one knows what all of the pieces are or realizes the limits of their own understanding.

Experts tend to exaggerate the centrality of their roles. Toxicologists may assume that everyone needs to know what they found when feeding rats a potential carcinogen or when testing groundwater near a landfill, even though additional information is always needed to make use of those results (e.g., physiological differences among species, routes of human exposure, compensating benefits of the exposure).

Another source of confusion is the failure of experts to remind laypeople of the acknowledged limits of the experts' craft. For example, cost-benefit analysts seldom remind readers that the calculations consider only total costs and benefits and, hence, ignore questions of who pays the costs and who pays the benefits (Bentkover et al., 1985; Smith and Desvousges, 1986).

Finally, environmental management is an evolving field that is only beginning to establish comprehensive training programs and methods, making it hard for anyone to know what the full picture is and how their work fits into it.

An enterprising journalist with a modicum of technical knowledge should be able to get specialists to tell their stories in fairly plain English and to cope with moderate evasiveness or manipulation. However, what is the journalist to do when the experts do not know what they do not know? One obvious solution is to talk to several experts with maximally diverse backgrounds. Yet, sometimes such a perfect mix is hard to find. Available experts can all have common limitations of perspective.

Another solution is to use a checklist of issues that need to be covered in any comprehensive environmental story. Scientists themselves use such lists to ensure that their own work is properly performed, documented, and reported. Such a protocol does not create knowledge for the expert any more than it would provide an education to the journalist. It does, however, help users exploit all they know—and acknowledge what they leave out.

Some protocols that can be used in looking at risk analyses are the causal model, the fault tree, a materials and energy flow diagram, and a risk analysis checklist.

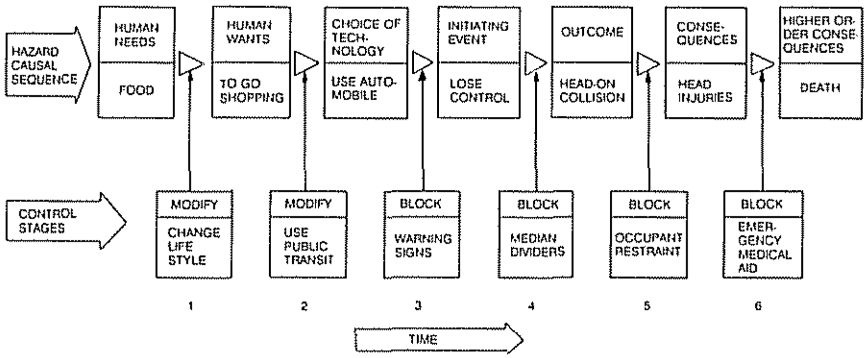

FIGURE II.1 The causal chain of hazard evolution. The top line indicates seven stages of hazard development, from the earliest (left) to the final stage (right). These stages are expressed generically in the top of each box and in terms of a sample motor vehicle accident in the bottom. The stages are linked by causal pathways denoted by triangles. Six control stages are linked to pathways between hazard states by vertical arrows. Each is described generically as well as by specific control actions. Thus control stage 2 would read: "You can modify technology choice by substituting public transit for automobile use and thus block the further evolution of the motor vehicle accident sequence arising out of automobile use." The time dimension refers to the ordering of a specific hazard sequence; it does not necessarily indicate the time scale of managerial action. Thus, from a managerial point of view, the occurrence of certain hazard consequences may trigger control actions that affect events earlier in the hazard sequence. SOURCE: Figure—Bick et al., 1979; caption—Fischhoff, Lichtenstein, et al., 1981.

The Causal Model

The causal model of hazard creation is a way to organize the full set of factors leading to and from an environmental mishap, both when getting the story and when telling it. The example in Figure II.1 is an automobile accident, traced from the need for transportation to the secondary consequence of the collision. Between each stage, there is some opportunity for an intervention to reduce the risk of an accident. By organizing information about the hazard in a chronological sequence, this scheme helps ensure that nothing is left out, such as the deep-seated causes of the mishap (to the left) and its long-range consequences (to the right).

Applied to an "irregular event" at a nuclear power station, for example, this protocol would work to remind a reporter of such (left-handed) causes as the need for energy and the need to protect the large capital investment in that industry and such (right-handed) consequences as the costs of retooling other plants designed like the

affected plant or the need to burn more fossil fuels if the plant is taken off line (without compensating reductions in energy consumption).

The Fault Tree

A variant on this procedure is the fault tree (Figure II.2), which lays out the sequence of events that must occur for a particular accident to happen (Green and Bourne, 1972; U.S. Nuclear Regulatory Commission, 1983). Actual fault trees, which can be vastly more involved than this example, are commonly used to organize the thinking and to coordinate the work of those designing complex technologies such as nuclear power facilities and chemical plants. At times, they are also used to estimate the overall riskiness of such facilities. However, the numbers produced are typically quite imprecise (U.S. Nuclear Regulatory Commission, 1978).

In effect, fault trees break open the right-handed parts of a causal model for detailed treatment. They can help a reporter to

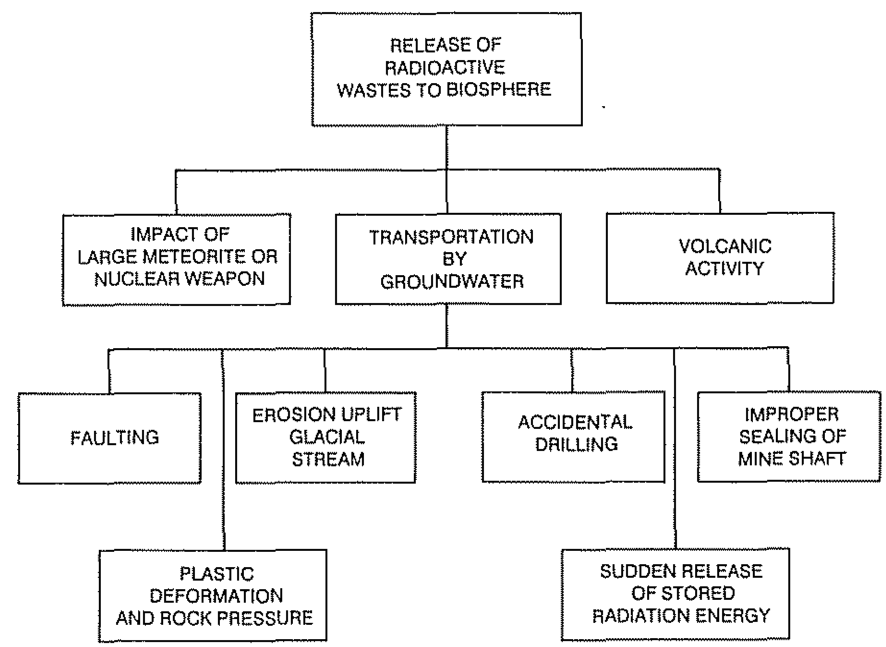

FIGURE II.2 Fault tree indicating the possible ways that radioactivity could be released from deposited wastes after the closure of a repository. SOURCE: Slovic and Fischhoff, 1983.

order the pieces of an accident story collected from different sources, see where an evolving incident (e.g., Three Mile Island or a leaking waste dump) is heading, and find out what safety measures were or were not taken.

Materials and Energy Flow Diagrams

The next model (Figure II.3) is adapted from the engineering notion of a materials or energy flow diagram. If something is neither created nor destroyed in a process, then one should be able to account schematically for every bit of it. In environmental affairs, one wants to account for all toxic materials. It is important to know where each toxic agent comes from and where each goes.

Keeping track of a substance can help anticipate where problems will appear, recur, and disappear. It can reveal when a problem has actually been treated and when it has merely been shifted to another time, place, or jurisdiction. With a story like EDB (ethylene dibromide, a fungicide used on grain) (Sharlin, 1987), such a chart would have encouraged questions such as, does it decay with storage or does it become something even worse when cooked and digested? Applying this approach led Harriss and Hohenemser (1978) to conclude that pollution controls had not reduced the total amount of mercury released into the environment, but only the distribution of releases (replacing a few big polluters with many smaller ones). In creating such figures, it is important to distinguish between where a substance is supposed to go and where it actually goes.

A comparable figure might be drawn to keep track of where the money goes, identifying the beneficiaries and losers resulting from different regulatory actions. With the EDB story, such a chart would have encouraged questions about who would eventually pay for the grain lost to pests if that chemical were not used. That is, would reducing the risk of EDB reduce producers' profits or increase consumers' prices? In the former case, failure to ban EDB looks much more callous than in the latter.

A Risk Analysis Checklist

The fourth aid (Figure II.4) is a list of questions that can be asked in a risk analysis (or of a risk analyst) in order to clarify what problem has been addressed and how well it has been solved.

This list was compiled for a citizens' group concerned with pesticides. Its members had mastered many substantive details of the

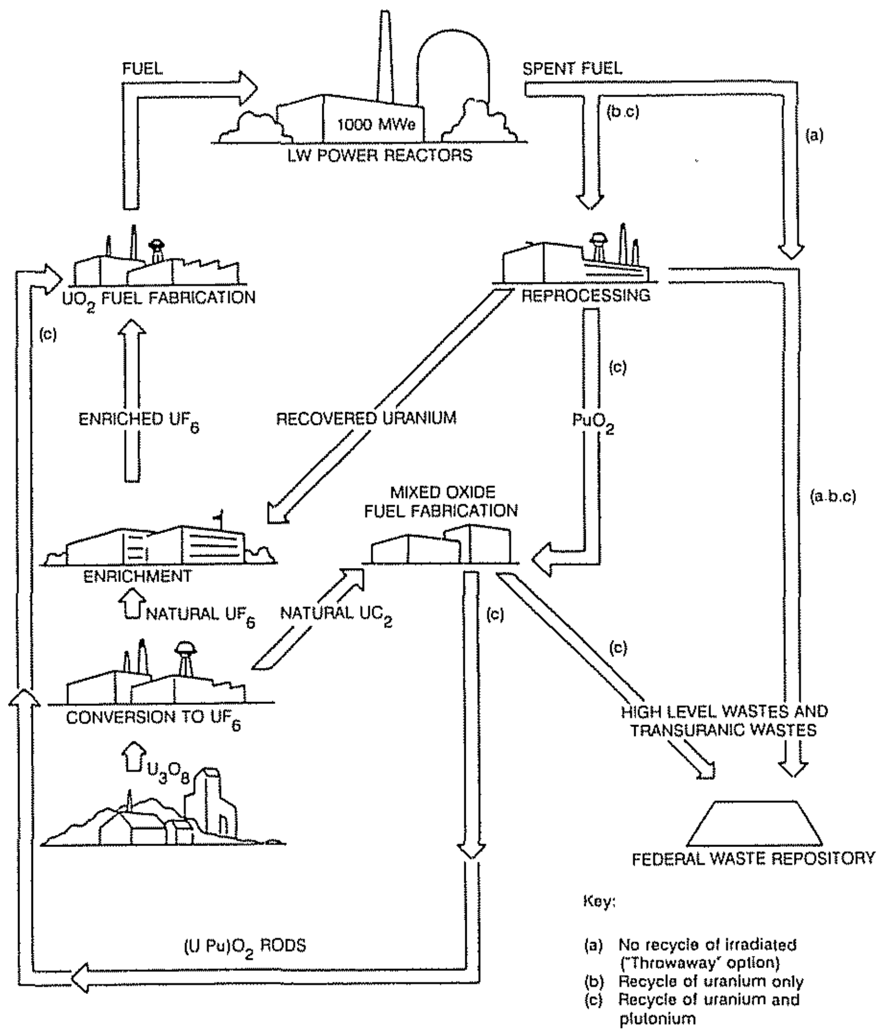

FIGURE II.3 Materials and energy flow diagram: Current options for the nuclear fuel cycle. SOURCE: Gotchy, 1983.

discipline, such as toxicology and biochemistry, involved in pesticide management, when suddenly they were confronted with a new procedure—risk analysis. In principle, risk analysis does no more than organize information from substantive disciplines in a way that allows overall estimates of risk to be computed. It can facilitate citizen access by forcing all the facts out on the table.

1. Does the risk analysis first state the health damage that may occur and then present the odds (i.e., the risk analysis)?
2. Is there enough information available on the factors that are most crucial to risk calculations?
3. If some of the data are missing, but there are enough to approach a risk assessment, are the missing data labeled as such?
4. Does the risk analysis disclose forthrightly the points at which it is based on guesswork?
5. Are various risk factors allowed to assume a variety of numbers depending on uncertainties in the data and/or various interpretations of the data?
6. Does the risk analysis multiply its probabilities by the number of people exposed to produce the number of people predicted to suffer damage?
7. Does the risk analysis disclose the confidence limits for its projections and the method of arriving at those confidence limits?
8. Are considerations of individual sensitivities, exposure to multiple pesticides, cumulative effects, and effects other than cancer, birth defects, and mutations included in the risk analysis?
9. Are all data and processes of the risk analysis open to public scrutiny?
10. Has an independent peer review of the risk analysis been funded and made public?
11. Are questions of (a) involuntary exposure, (b) who bears the risks and who reaps the benefits, and (c) alternatives to pesticide use being considered alongside the risk analysis?
12. Are alternatives to pesticide use also being extensively analyzed for risk or lack of risk?
13. Are the processes of risk analysis and risk policy separate?

FIGURE II.4 Risk analysis checklist. SOURCE: Northwest Coalition for Alternatives to Pesticides, 1985.

However, unless one can penetrate all its formalisms, risk analysis can mystify and obscure the facts rather than reveal them. Such a checklist can clarify what an analysis has done in terms approximating plain English.

WHAT IS THE HARD SCIENCE RELATED TO THE PROBLEM?

With most "interesting" hazards, the data run out long before enough is known to estimate their risks and benefits as precisely as one would want. Much of risk management involves going beyond the available data either to guess at what the facts might be or to figure out how to live with uncertainty. Obviously, one wants to reduce this uncertainty by making the best of the hard data available.

Unfortunately, there is no short-cut to providing observers with ways to read critically all of the kinds of science that could be invoked in the course of characterizing a risk. There are too many sciences to consider and too many nuances in each type of science to know

about in assessing the validity of studies conducted in any one field. Even the social sciences, which seem relatively accessible (compared with the physical sciences) and the results of which can be rendered into common English, routinely foil the efforts of amateur scientists.

These failures can be seen most clearly in the attempts by non-social scientists to make factual statements about the behavior of laypeople, solely on the basis of their untrained anecdotal observations. Such speculations can mislead more than inform if they are made without realizing that they lack the discipline of science.

The complexities of science arise in the details of creating, analyzing, and interpreting specific sets of data. To give a feeling for these strengths and limits of scientific research, several examples drawn from social science research into risk perception and communication are presented here. Each science has its own nuances. Featuring this science also provides background for interpreting the social science results described below.

Like speculations about chemical reactions, speculations about human behavior must be disciplined by fact. Such speculations make important statements about people and their capabilities, and failure to validate them may mean arrogating to oneself considerable political power. Such happens, for example, when one says that people are so poorly informed (and ineducable) they require paternalistic institutions to defend them, and, furthermore, they might be better off surrendering some political rights to technical experts. It also happens, at the other extreme, when one claims that people are so well informed (and offered such freedom of choice) one need not ask them anything at all about their desires; to know what they want, one need only observe their behavior in the marketplace. It also happens when we assume that people are consummate hedonists, rational to the extreme in their consumer behavior but totally uncomprehending of broader economic issues, so we can impose effective fiscal policies on them without being second-guessed.

One reason for the survival of such simplistic and contradictory positions is political convenience. Some people want the lay public to participate actively in hazard management decisions, and need to be able to describe the public as competent; others need an incompetent public to legitimate an expert elite. A second reason is theoretical convenience. It is hard to build models of people who are sometimes wise and sometimes foolish, sometimes risk seeking and sometimes risk averse. A third reason is that one can effortlessly speculate about human nature and even produce a bit of supporting anecdotal information. Indeed, good social theory may be so rare because poor social theory is so easy.

Judgments of Risk

At first sight, assessing the public's risk perceptions would seem to be very straightforward. Just ask questions like, "What is the probability of a nuclear core meltdown?" or "How many people die annually from asbestos-related diseases?" or "How does wearing a seat belt affect your probability of living through the year?" Once the results are in, they can be compared with the best available technical estimates, with deviations interpreted as evidence of respondents' ignorance.

Unfortunately, how one asks the question may in large part determine the content (and apparent wisdom) of the response. Lichtenstein and her colleagues (Lichtenstein et al., 1978) asked two groups of educated laypeople to estimate the frequency of death in the United States from each of 40 different causes. The groups differed only in the information that was given to them about one cause of death in order to help scale their responses. One group was told about 50,000 people die annually in motor vehicle accidents, and the other was told about 1,000 annual deaths result from electrocution. Both reports were accurate, but receiving a larger number increased the estimates of most frequencies for respondents in the motor vehicle accident group. This is a special case of a general psychological phenomenon called "anchoring," whereby people's responses are pulled toward readily available numbers in cases in which they do not know exactly what to say (Poulton, 1968, 1977; Tversky and Kahneman, 1974). Such anchoring on the original number changed the smallest estimates by roughly a factor of 5.

Fischhoff and MacGregor (1983) asked people to judge the lethality of various potential causes of death using one of four formally equivalent formats (e.g., "For each afflicted person who dies, how many survive?" or "For each 100,000 people afflicted, how many will die?"). Table II.1 expresses their judgments in a common format and reveals even more dramatic effects of question phrasing on expressed risk perceptions. For example, when people estimated the lethality rate for influenza directly (column 1), their mean response was 393 deaths per 100,000 cases. When told that 80 million people catch influenza in a normal year and asked to estimate the

TABLE II.1 Lethality Judgments with Four Different Response Modes (geometric mean)

	Death Rate Per 100,000 Afflicted				
Condition	Estimated Lethality Rate	Estimated Number Who Die	Estimated Survival Rate	Estimated Number Who Survive	Actual Lethality Rate
Influenza	393	6	26	511	1
Mumps	44	114	19	4	12
Asthma	155	12	14	599	33
Venereal disease	91	63	8	111	50
High blood pressure	535	89	17	538	76
Bronchitis	162	19	43	2111	85
Pregnancy	67	24	13	787	250
Diabetes	487	101	52	5666	800
Tuberculosis	852	1783	188	8520	1535
Automobile accidents	6195	3272	31	6813	2500
Strokes	11,011	4648	181	24,758	11,765
Heart attacks	13,011	3666	131	27,477	16,250
Cancer	10,889	10,475	160	21,749	37,500

NOTE: The four experimental groups were given the following instructions:
(a) Estimate lethality rate: For each 100,000 people afflicted, how many die?
(b) Estimate number who die: X people were afflicted, how many died?
(c) Estimate survival rate: For each person who died, how many were afflicted but survived?
(d) Estimate number who survive: Y people died, how many were afflicted but did not die?
Responses to (b), (c), and (d) were converted to deaths per 100,000 to facilitate comparisons.

SOURCE: Fischhoff and MacGregor, 1983.

number who die (column 2), their mean response was 4800, representing a death rate of only 6 per 100,000 cases. This slight change in the question changed the estimated rate by a factor of more than 60. Similar discrepancies occurred with other questions and other hazards. One consequence for risk communicators is that whether laypeople intuitively overestimate or underestimate risks (or perceive them accurately) depends on what question they are asked.

In a recent study at an Ivy League college (Linville et al., 1988), students were asked to give estimates of the probability that the AIDS virus could be transmitted from a man to a woman in a single case of unprotected sex. The median estimate was about 10 percent, considerably above current scientific estimates (Fineberg, 1988). However, when asked to give estimates for the probability of transmission in 100 cases of unprotected sex, the median answer was about 25 percent. This risk estimate is considerably more in line with scientific thinking—so that an investigator asking this question would have a considerably more optimistic assessment of the state of public understanding. Unfortunately, it is also completely inconsistent with the single-case estimates produced by the same individuals. If one believes in a single-case probability of 10 percent, then transmission should be a virtual certainty with 100 exposures. Such failure to see how small risks mount up over repeated exposures has been observed in such diverse settings as the risks from playing simple gambles (Bar-Hillel, 1973), driving (Slovic et al., 1978), and relying on various contraceptive devices (Shaklee et al., 1988).

Such effects are hardly new; indeed, some have been recognized for close to 100 years. Early psychologists discovered that different numerical judgments may be attached to the same physical stimulus (e.g., the loudness of a tone) as a function of whether the set of alternatives is homogeneous or diverse, and whether the respondent makes one or many judgments. Even when the same presentation is used, different judgments might be obtained with a numerical or a comparative (ordinal) response mode, with instructions stressing speed or accuracy, with a bounded or an unbounded response set, and with verbal or numerical response labels.

The range of these effects may suggest that the study of judgment is not just difficult, but actually impossible. Closer inspection, however, reveals considerable orderliness underlying this apparent chaos (Atkinson et al., 1988; Carterette and Friedman, 1974; Woodworth and Schlosberg, 1954).

Judgments of Values

Once the facts of an issue have been estimated and communicated, it is usually held that laypeople should (in a democracy) be asked about their values. What do they want—after the experts have told them what they can (conceivably) have? Here, too, the straightforward strategy of "just ask them" runs into trouble.

The problem of poorly (or even misleadingly) worded questions in attitude surveys is well known, although not necessarily well resolved (Bradburn and Sudman, 1979; National Research Council, 1982; Payne, 1952; Zeisel, 1980). For example, a major trade publication (Ventner, 1979) presented the results of a survey of public attitudes toward the chemical industry containing the following question:

> Some people say that the prime responsibility for reducing exposure of workers to dangerous substances rests with the workers themselves, and that all substances in the workplace should be clearly labeled as to their levels of danger and workers then encouraged or forced to be careful with these substances. Do you agree or disagree?

It is hard to know what one is endorsing when one says "Yes," "No," or "I don't know" to such a complex and unclear question.

Although annoying, ambiguous wording is, in principle, a relatively easy problem to deal with because there are accepted ways to "do it right." Much more complicated are cases in which seemingly arbitrary aspects of how a question is posed affect the values. Parducci (1974) has found that judged satisfaction with one's state in life may depend on the range of possible states mentioned in the question put to people. In an attempt to establish a dollar value for aesthetic degradation of the environment, Brookshire et al. (1976) asked visitors to Lake Powell how much they would be willing to pay in increased users' fees in order not to have an ugly (coal-fired) power plant looming on the opposite shore. They asked "Would you pay $1, $2, $3?" and so on, until the respondent answered "No" and then they retreated in decrements of a quarter (e.g., "Would you pay $5.75, $5.50, . . .?"). Rather different numerical values might have been obtained had the bidding procedure begun at $100 and decreased by steps of $10 or with other plausible variants. Any respondents who were not sure what they wanted in dollars and cents might naturally and necessarily look to the range of options presented, the difference between first and second options, and so on, for cues as to what are reasonable and plausible responses (Cummings et al., 1986; Smith and Desvousges, 1986).

At first glance, it might seem as though questions of value are the last redoubt of unaided intuition. Who knows better than an individual what he or she prefers? When people are considering simple, familiar events with which they have direct experience, it may be reasonable to assume that they have well-articulated opinions. Regarding the novel, global consequences potentially associated with CO_2-induced climatic change, nuclear meltdowns, or genetic engineering, that may not be the case. Our values may be incoherent, not thought through. In thinking about what are acceptable levels of risk, for example, we may be unfamiliar with the terms in which issues are formulated (e.g., social discount rates, minuscule probabilities, or megadeaths). We may have contradictory values (e.g., a strong aversion to catastrophic losses of life and a realization that we are no more moved by a plane crash with 500 fatalities than by one with 300). We may occupy different roles in life (parents, workers, children) that produce clear-cut but inconsistent values. We may vacillate between incompatible, but strongly held, positions (e.g., freedom of speech is inviolate, but should be denied to authoritarian movements). We may not even know how to begin thinking about some issues (e.g., the appropriate trade-off between the opportunity to dye one's hair and a vague, minute increase in the probability of cancer 20 years from now). Our views may undergo changes over time (say, as we near the hour of decision or of experiencing the consequence) and we may not know which view should form the basis of our decision.

An extreme, but not uncommon, situation is having no opinion and not realizing it. In that state, we may respond with the first thing that comes to mind once a question is asked and then commit ourselves to maintaining that first expression and to mustering support for it, while suppressing other views and uncertainties. As a result, we may be stuck with stereotypical or associative responses, generated without serious contemplation.

Once an issue has been evoked, it must be given a label. In a world with few hard evaluative standards, such symbolic interpretations may be very important. While the facts of abortion remain constant, individuals may vacillate in their attitude as they attach and detach the label of murder. Figure II.5 shows two versions of the same gamble, differing only in whether one consequence is labeled a "sure loss" or an "insurance premium." Most people dislike the former and like the latter. When these two versions are presented sequentially, people often reverse their preferences for the two options (Hershey and Shoemaker, 1980). Figure II.6 shows a labeling effect that produced a reversal of preference with practicing physicians; most preferred treatment A over treatment B, and treatment D over treatment C, despite the formal equivalence of A and C and of B and D. Saving lives and losing lives afforded very different perspectives on the same problem.

People solve problems, including the determination of their own values, with what comes to mind. The more detailed, exacting, and creative their inferential process, the more likely they are to think of all they know about a question. The briefer that process becomes,

Insurance

Imagine that you must play a gamble in which you can lose but cannot win Specifically, this gamble exposes you to:

1 chance in 4 to lose £200
(and 3 chances in 4 to lose nothing).

You can either take a chance with the gamble or insure against the £200 loss by buying a policy for a premium of £50. If you buy this insurance, you cannot lose £200, but you must pay the £50 premium.

Please indicate what you would do in this situation.

Preference

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: 1 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 would probably prefer not to be in either of these situations, but, if forced either to play the gamble (A) or to accept the certain loss (B), which would you prefer to do?

FIGURE II.5 Two formulations of a choice problem: insurance versus certain loss. SOURCE: Fischhoff et al., 1980.

the more they will be controlled by the relative accessibility of various considerations. Accessibility may be related to importance, but it is also related to the associations that are evoked, the order in which questions are posed, imaginability, concreteness, and other factors only loosely related to importance. As one example of how an elicitor may (perhaps inadvertently) control respondents' perspective, Turner (1980) observed a large difference in responses to a simple question such as "Are you happy?" on two simultaneous surveys of the same population (Figure II.7). The apparent source of the difference was that one (NORC) preceded the happiness question with a set of questions about married life. In the United States, married people are generally happier than unmarried people. Reminding them of that aspect of their life apparently changed the information that they brought to the happiness question.

It would be comforting to be able to say which way of phrasing these questions is most appropriate. However, there is no general answer. One needs to know why the question is being asked (Fischhoff

Lives Saved

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. The accepted scientific estimate of the consequences of the program are as follows:

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved.

Which of the two programs would you favor?

Lives Lost

If Program C is adopted, 400 people will die.

If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

Which of the two programs would you favor?

FIGURE II.6 Two formulations of a choice problem: lives saved versus lives lost. SOURCE: Tversky and Kahneman, 1981. Copyright © 1981 by the American Association for the Advancement of Science.

and Furby, 1988). If one wants to predict the quality of casual encounters, then a superficial measure of happiness may suffice. However, an appraisal of national malaise or suicide potential may require a questioning procedure that evokes an appreciation of all components of respondents' lives. It has been known for some time that white interviewers evoke more moderate responses from blacks on race-related questions than do black interviewers. The usual response has been to match the races of interviewer and interviewee (Martin, 1980). This solution may be appropriate for predicting voting behavior or conversation in same-race bars, but not for predicting behavior of blacks in white-dominated workplaces.

The fact that one has a question is no guarantee that respondents have answers, or even that they have devoted any prior thought to the matter. When one must have an answer (say, because public input is statutorily required), there may be no substitute for an elicitation procedure that educates respondents about how they might look at the question. The possibilities for manipulation in such interviews are obvious. However, one cannot claim to be serving respondents' best interests (letting them speak their minds) by asking a question

that only touches one facet of a complex and incompletely formulated set of views.

Refining Common Sense

Social scientists often find themselves in a no-win situation. If they describe their work in technical jargon, no one wants to listen. If they use plain language, no one feels a need to listen. Listeners feel that they "knew it all along" and that the social scientist was just "affirming the obvious" or "validating common sense." One possible antidote to this feeling is to point out the evidence showing that, in hindsight, people exaggerate how much they could have known in foresight, leading them to discount the informativeness of scientific

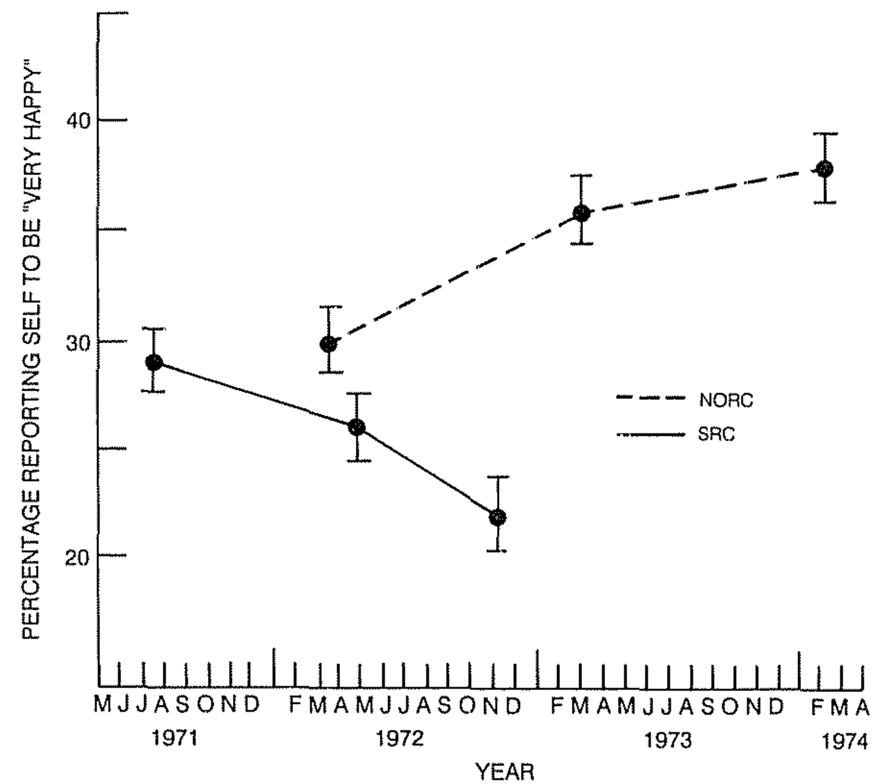

FIGURE II.7 Trends in self-reported happiness derived from sample surveys of the noninstitutionalized population of the continental United States aged 18 and over. Error bars demark ±1 standard error around sample estimate. SOURCE: Turner, 1980.

reports (Slovic and Fischhoff, 1977). A second antidote is to note that common sense often makes contradictory predictions (e.g., two heads are better than one versus too many cooks spoil the broth; absence makes the heart grow fonder versus out of sight, out of mind). Research is needed to determine which version of common sense is correct or what their respective ranges of validity are. A third strategy, adopted immediately below, is to present empirical results that contradict conventional wisdom (Lazarsfeld, 1949).

Informing People About Risks

It is often claimed that people do not want to know very much about the health risks they face, since such information makes them anxious. Moreover, they cannot use that information very productively, even if it is given. If true, these claims would make it legitimate for someone else (e.g., physicians, manufacturers, government) to decide what health (and therapeutic) risks are acceptable, and not to invest too much effort on information programs. A number of investigators, however, have replaced anecdotal evidence with systematic observation and have found that, by and large, people want to be told about potential risks (Alfidi, 1971; Weinstein, 1980a). In clinical settings, this desire has been observed with such risky practices as psychotropic medication (Schwarz, 1978), endoscopy (Roling et al., 1977), and oral contraceptives (Applied Management Sciences, 1978; Joubert and Lasagna, 1975). Figure II.8 shows respondents' strong opinions about the appropriate use of a pamphlet designed to explain the risks faced by temporary workers in a nuclear power plant. Ninety percent of these individuals gave the most affirmative answer possible to the question, "If you had taken such a job without being shown this pamphlet, would you feel that you had been deprived of necessary information?" (Fischhoff, 1981).

Risk-Taking Propensity

We all know that some people are risk takers and others are risk avoiders; some are cautious, whereas others are rash. Indeed, attitude toward risk might be one of the first attributes that comes to mind when one is asked to describe someone else's personality. In 1962, Slovic compared the scores of 82 individuals on nine different measures of risk taking. He found no consistency at all in people's propensity for taking risks in the settings created by the various tests (Slovic, 1962). Correlations ranged from −.35 to .34, with a mean of

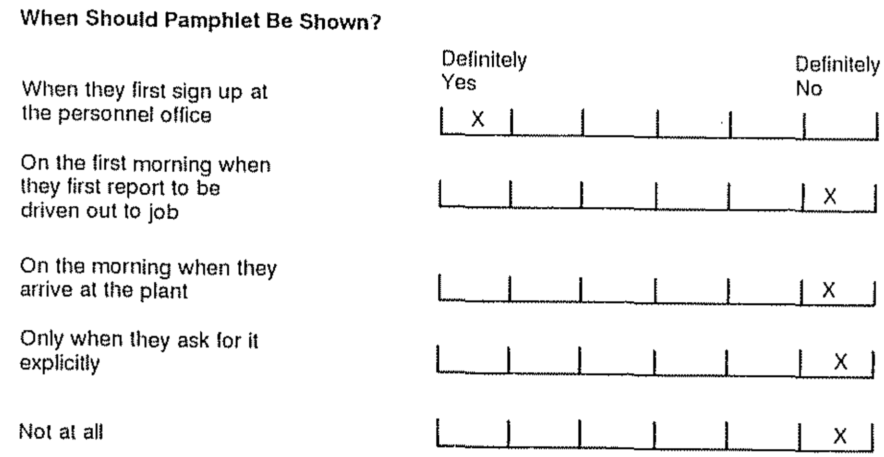

FIGURE II.8 Opinions about the appropriate use of a pamphlet describing the risks associated with temporary work in a facility handling nuclear materials. Respondents were drawn from the readers of a student newspaper and from unemployed individuals at a state labor exchange. The "X" on each line represents the mean response to a question by the 173 individuals. SOURCE: Fischhoff, 1981.

.006. That is, people who are daring in one context may be timid in another, a result that has been replicated in numerous other studies (Davidshofer, 1976).

The surprising nature of these results may tell us something about ourselves as well as about the people we observe. One of the most robust psychological discoveries of the past 20 years has been identification of the *fundamental attribution error,* the tendency to view ourselves as highly sensitive to the demands of varying situations, but to see others as driven to consistent behavior by dominating personality traits (Nisbett and Ross, 1980). This misperception may be attributable to the fact that we typically see most others in only one role, as workers or spouses or parents or tennis players or drivers or whatever, in which the situational pressures are quite consistent. Thus, we may observe accurately the evidence available to us, but fail to understand the universe from which these data are drawn.

Protective Behavior

For years, the United States has been building flood control projects. Despite these great expenditures, flood losses today (in constant dollars) are greater than they were before this enterprise began. Apparently, the behavioral models of the dam and levee builders failed to account for the extent to which eliminating the recurrence of small-to-moderate floods reduced residents' (and particularly newcomers') sensitivity to flood dangers, which in turn led to overbuilding the flood plain. As a result, when the big floods come (about once every 100 years), exceeding the containment capacity of the protective structures, much more lies in their path (White, 1974).

The official response to this situation has been the National Flood Insurance Program (Kunreuther et al., 1978), designed according to economic models of human behavior, which assumes that flood plain residents are all-knowing, all-caring, and entirely "rational" (as defined by economics). Initially, premiums were greatly subsidized by the federal government to make the insurance highly attractive; these subsidies were to be withdrawn gradually once the insurance-buying habit was established. Unfortunately for the program, few people bought the insurance. The typical explanation for this failure was that residents expected the government to bail them out in the event of flood. However, a field survey found this speculation, too, to be in error. Flood plain residents reported that they expected no help, feeling that they were willingly bearing an acceptable risk. When residents thought about insurance at all, they seemed to rely on a melange of ad hoc principles like, "I can't worry about everything" and "The chances of getting a return (reimbursement) on my investment (premium) are too small," rather than on the concepts and procedures of economics (Kunreuther et al., 1978; Slovic et al., 1977).

ADHERENCE TO ESSENTIAL RULES OF SCIENCE

Looking hard at other sciences would reveal them to be similarly complicated, and similarly surprising. Sciences may not reveal their intricacies readily, but committed citizen activists have often proven themselves capable of mastering enough of the relevant science to be able to ask hard questions about risk issues that interest them (Figure II.4, for example, was created as a step toward this end). Many, of course, do not, and none could learn the hard questions about all of the sciences impinging on complex risk issues. This is, however, an option for those who care enough.

Short of such intense involvement, it is possible to ask some

generic questions about almost any science. These are ways of asking "How good could it be?", given the conditions of its production.

Perhaps the most basic question that one can ask about any bit of science introduced into an environmental dispute, whether it be a single rodent bioassay or a full-blown risk analysis, is whether it actually represents a bit of science. In applied settings, one often finds evidence that fails to adhere to such essential rules of science as: (1) subjecting the study to critical peer review; (2) making all data available to other investigators; (3) evaluating the statistical reliability of results; (4) considering alternative explanations of the results; (5) relating new results to those already in the literature; and (6) pointing out critical assumptions that have not been empirically verified. Studies that fail to follow such procedures may be attempting to assume the rights, but not the responsibilities of science. Conversely, good science can come even from partisan sources (e.g., industry labs, environmental activists), if the rules are followed.

The definitiveness of science is bounded not only by the process by which it is conducted, but also by the object of its study. Some topics are simply easier than others, allowing for results clouded by relatively little uncertainty. Unfortunately for the rapid understanding and resolution of problems, risk management often demands understanding of inherently difficult topics.

This difficulty for risk managers can be seen as a by-product of one fortunate feature of the natural environment, namely, that the most fearsome events are quite infrequent. Major floods, disastrous plagues, and catastrophic tremors are all the exception rather than the rule. Social institutions attempt to constrain hazards of human origin so that the probability of their leading to disaster is low. However great their promised benefit, projects that might frequently kill large numbers of people are unlikely to be developed. The difficult cases are those in which the probability of a disaster is known to be low, but we do not know just how low. Unfortunately, quantitative assessment of very small probabilities is often very difficult (Fairley, 1977).

At times, one can identify a historical record that provides frequency estimates for an event related to the calamity in question. The U.S. Geological Survey has perhaps 75 years of reliable data on which to base assessments of the likelihood of large earthquakes (Burton et al., 1978). Iceland's copious observations of ice-pack movements over the last millennium provide a clue to the probability of an extremely cold year in the future (Ingram et al., 1978). The absence of a full-scale meltdown in 500 to 1000 reactor-years of nuclear power plant operation sets some bounds on the probability of future meltdowns (Weinberg, 1979). Of course, extrapolation from any of these historical records is a matter of judgment. The great depth and volume of artificial reservoirs may enhance the probability of earthquakes in some areas. Increased carbon dioxide concentrations in the atmosphere may change the earth's climate in ways that amplify or moderate yearly temperature fluctuations. Changes in design, staffing, and regulation may render the next 1000 reactor-years appreciably different from their predecessors. Indeed, any attempt to learn from experience and make a technology safer renders that experience less relevant for predicting future performance.

Even when experts agree on the interpretation of records, a sample of 1000 reactor-years or calendar-years may be insufficient. If one believes the worst-case scenarios of some opponents of nuclear power, a 0.0001 chance of a meltdown (per reactor-year) might seem unconscionable. However, we will be into the next century before we will have enough on-line experience to know with great confidence whether the historical probability is really that low.

HOW DOES JUDGMENT AFFECT THE RISK ESTIMATION PROCESS?

To the extent that historical records (or records of related systems) are unavailable, one must rely on conjecture. The more sophisticated conjectures are based on models such as the fault-tree and event-tree analyses of a loss-of-coolant accident upon which the Reactor Safety Study was based (U.S. Nuclear Regulatory Commission, 1975). As noted in Figure II.2, a fault tree consists of a logical structuring of what would have to happen for an accident (e.g., a meltdown) to occur. If sufficiently detailed, it will reach a level of specificity for which one has direct experience (e.g., the operation of individual valves). The overall probability of system failure is determined by combining the probabilities of the necessary component failures.

The trustworthiness of such an analysis hinges on the experts' ability to enumerate all major pathways to disaster and on the assumptions that underlie the modeling effort. Unfortunately, a modicum of systematic data and many anecdotal reports suggest that experts may be prone to certain kinds of errors and omissions. Table

TABLE II.2 Some Problems in Structuring Risk Assessments

Failure to consider the ways in which human errors can affect technological systems.

Example: Owing to inadequate training and control room design, operators at Three Mile Island repeatedly misdiagnosed the problems of the reactor and took inappropriate actions (Sheridan, 1980; U.S. Government, 1979)

Overconfidence in current scientific knowledge

Example: DDT came into widespread and uncontrolled use before scientists had even considered the possibility of the side effects that today make it look like a mixed, and irreversible, blessing (Dunlap, 1978).

Failure to appreciate how technological systems function as a whole.

Example: The DC-10 failed in several early flights because its designers had not realized that decompression of the cargo compartment would destroy vital control systems (Hohenemser, 1975).

Slowness in detecting chronic, cumulative effects.

Example: Although accidents to coal miners have long been recognized as one cost of operating fossil-fueled plants, the effects of acid rain on ecosystems were slow to be discovered (Rosencranz and Wetstone, 1980).

Failure to anticipate human response to safety measures.

Example: The partial protection afforded by dams and levees gives people a false sense of security and promotes development of the flood plain. Thus, although floods are rarer, damage per flood is so much greater that the average yearly loss in dollars is larger than before the dams were built (Burton et al., 1978).

Failure to anticipate common-mode failures, which simultaneously afflict systems that are designed to be independent.

Example: Because electrical cables controlling the multiple safety systems of the reactor at Browns Ferry, Alabama, were not spatially separated, all five emergency core-cooling systems were damaged by a single fire (Jennergren and Keeney, 1982; U.S. Government, 1975)

SOURCE: Fischhoff, Lichtenstein, et al., 1981a.

II.2 suggests some problems that might underlie the confident veneer of a formal model.

When the logical structure of a system cannot be described to allow computation of its failure probabilities (e.g., when there are large numbers of interacting systems), physical or computerized simulation models may be used. If one believes the inputs and the programmed interconnections, one should trust the results. What happens, however, when the results of a simulation are counterintuitive or politically awkward? There may be a strong temptation to try it again, adjusting the parameters or assumptions a bit, given that many of these are not known with certainty in the first place. Susceptibility to this temptation could lead to a systematic and subtle bias in modeling. At the extreme, models would be accepted only if they confirmed expectations.

Acknowledging the Role of Judgment

Although the substance of sciences differs greatly, sciences do have in common the fact that they are produced by the minds of mortals. Those minds may contain quite different facts, depending on the disciplines in which they were trained. However, it is reasonable to suppose that they operate according to similar principles when they are pressed to make speculations—taking them beyond the limits of hard data—in order to produce the sorts of assessments needed to guide risk managers.

Indeed, the need for judgment is a defining characteristic of risk assessment (*Federal Register* 49(100):21594–21661). Some judgment is, of course, a part of all science. However, the policy questions that hinge on the results of risk assessments typically demand greater scope and precision than can be provided by the "hard" knowledge that any scientific discipline currently possesses. As a result, risk assessors must fill the gaps as best they can. The judgments incorporated in risk assessments are typically those of esteemed technical experts, but they are judgments nonetheless, taking one beyond the realm of established fact and into the realm of educated opinions that cannot immediately be validated.

Judgment arises whenever materials scientists estimate the failure rates for valves subjected to novel conditions (Joksimovich, 1984; Östberg et al., 1977), whenever accident analysts attempt to recreate operators' perceptions of their situation prior to fatal mishaps (Kadlec, 1984; Pew et al., 1982), when toxicologists choose and weight extrapolation models (Rodricks and Tardiff, 1984; Tockman and Lilienfeld, 1984), when epidemiologists assess the reasons for nonresponse in a survey (Joksimovich, 1984; National Research Council, 1982), when pharmacokineticists consider how consumers alter the chemical composition of foods (e.g., by cooking and storage practices) before they consume them (National Research Council, 1983a; O'Flaherty, 1984), when physiologists assess the selection bias in the individuals who volunteer for their experiments (Hackney and

Linn, 1984; Rosenthal and Rosnow, 1969), when geologists consider how the construction of underground storage facilities might change the structure of the rock media and the flow of fluids through them (Sioshansi, 1983; Travis, 1984), and when psychologists wonder how the dynamics of a particular group of interacting experts affect the distribution of their responses (Brown, 1965; Davis, 1969; Hirokawa and Poole, 1986).

The process by which judgments are produced may be as varied as the topics they treat. Individual scientists may probe their own experience for clues to the missing facts. Reviewers may be sponsored to derive the best conclusions that the literature can provide. Panels of specialists may be convened to produce a collective best guess. Trained interviewers may use structured elicitation techniques to extract knowledge from others. The experts producing these judgments may be substantive experts in almost any area of science and engineering, risk assessment generalists who take it upon themselves to extrapolate from others' work, or laypeople who happen to know more than anyone else about particular facts (e.g., workers assessing how respirators are really used, civil defense officials predicting how evacuation plans will work).

Few experts would deny that they do not know all the answers. However, detailed treatments of the judgments they make in the absence of firm evidence are seldom forthcoming (*Federal Register* 49(100):21594–21661). There appear to be several possible causes for this neglect. Knowing which is at work in a particular risk assessment establishes what effect, if any, the informal treatment of judgment has had.

One common reason for treating the role of judgment lightly is the feeling that everyone knows that it is there, hence there is no point in repeating the obvious. Although this feeling is often justified, acting on it can have two deleterious consequences. One is that all consumers of an assessment may not share the same feeling. Some of these consumers may not realize that judgment is involved, whereas others may suspect that the judgments are being hidden for some ulterior purpose. The second problem is that failure to take this step precludes taking the subsequent steps of characterizing, improving, and evaluating the judgments involved.

A second, complementary reason for doing little about judgment is the belief that nothing much can be done, beyond a good-faith effort to think as hard as one can. Considering the cursory treatment of judgmental issues in most methodological primers for risk analysts, this perception is understandable. Considering the importance of doing something and the extensive research regarding what can be done, it is, however, not justifiable. Although the research is unfamiliar to most practicing analysts, the study and cultivation of judgment have proven tractable. The vulnerability of analyses to judgmental difficulties means that those who ignore judgment for this reason may miss a significant opportunity to perform at the state of the art.

A third reason for ignoring judgment is being rewarded for doing so. At times, analysts discern some strategic advantage to exaggerating the definitiveness of their work. At times, analysts feel that they must make a begrudging concession to the demands of political processes that attend only to those who speak with (unjustifiable) authority. At times, the neglect of judgment is (almost) a condition of employment, as when employers, hearings officials, or contracting agencies require statements of fact, not opinion.

Diagnosing the Role of Judgment

The first step in dealing with the judgmental aspects of risk assessments is identifying them. All risk assessment, and most contemporary science, can be construed as the construction of models. These include both procedures used to assess discrete hazards (e.g., accidents), such as probabilistic risk analysis, and procedures used to assess continuous hazards (e.g., toxicity), such as dose-response curves or structural-activity relationships. Although these models take many forms, all require a similar set of judgmental skills, which can be used as a framework for diagnosing where judgment enters into analyses (and, subsequently, how good it is and what can be done about it). These skills are:

1. *Identifying the active elements of the hazardous system being studied.* These may be the physical components of a nuclear power plant (e.g., the valves, controls, and piping) (U.S. Nuclear Regulatory Commission, 1983), the environmental factors affecting the dispersal of toxins from a waste disposal site (e.g., geologic structure, rainfall patterns, adjacent construction) (Pinder, 1984), or the potential predictors of cancer in an epidemiological study (Tockman and Lilienfeld, 1984).

2. *Characterizing the interrelationships among these elements.* Not everything is connected to everything else. Reducing the set of interconnections renders the model more tractable, its results more comprehensible, and its data demands more manageable. The probabilistic risk analyst must judge which malfunctions in System X need to be considered when studying the performance of System Y. The epidemiologist needs to judge which interaction terms to include in regression models.

3. *Assessing the value of model parameters.* The amount of this kind of judgment varies greatly both across and within analyses. Some values have a sound statistical base (e.g., the number of chemical workers, as revealed by a decennial census), whereas others must be created from whole cloth (e.g., the sabotage rate at an as-yet-unconstructed plant 10 years in the future). Yet even the firmest statistics require some interpretation, for example, to correct for sampling and reporting biases or to adjust for subsequent changes in conditions.

4. *Evaluating the quality of the analysis.* Every analysis requires some summary statement of how good it is, whether for communicating its results to policymakers or for deciding whether to work on it more. Such evaluation requires consideration of both the substance and the purpose of the analysis. In both basic and applied sciences, the answer to "is the assessment good enough?" presupposes an answer to "good enough for what?"

5. *Adopting appropriate judgmental techniques.* Just as each stage in risk assessment requires different judgmental skills, it also requires different elicitation procedures. The reason for this is that each kind of information is organized in people's minds in a different way, and needs, therefore, to be extracted in a different way. For example, listing all possible mistakes that operators of a process-control industry might make is different than estimating how frequently each mistake will be made. The former requires heavy reliance on memory for instances of past errors, whereas the latter requires aggregation across diverse experiences and their extrapolation to future situations. Different experts (e.g., veteran operators, human factors theorists) may be more accustomed to thinking about the topic in one way rather than the other. Although transfer of information between these modes of thinking is possible, it may be far from trivial (Lachman et al., 1979; Tulving, 1972).

As noted earlier, studies with laypeople have found that seemingly subtle variations in how judgments are elicited can have large effects on the beliefs that are apparently revealed. These effects are most pronounced when people are least certain about how to respond, either because they do not know the answers or because they are unaccustomed to expressing themselves in the required terms. Thus, in extrapolating these results one must ask how expert the respondents are both in the topic requiring judgment and in using that response mode.

Assessing the Quality of the Judgment

If analysts have addressed the preceding steps conscientiously and left an audit trail of their work, all that remains is to review the protocol of the analysis to determine how heavily its conclusions depend on judgment and how adequate those judgments are likely to be. That evaluation should consider both the elicitation methods used and the judgmental capabilities of the experts. Ideally, the methods would have been empirically tested to show that they are: (1) compatible with the experts' mental representation of the problem, and (2) able to help the experts use their minds more effectively by overcoming common judgmental difficulties. Ideally, the experts would not only be knowledgeable about the topic, but also capable of translating that knowledge into the required judgments. The surest guarantees of that capability are having been trained in judgment or having provided judgments in conditions conducive to skill acquisition (e.g., prompt feedback).

How Good Are Expert Judgments?

As one might expect, considerably more is known about the judgmental processes of laypeople than about the judgmental processes of experts performing tasks in their areas of expertise. It is simply much easier to gain access to laypeople and create tasks about everyday events. Nonetheless, there are some studies of experts per se. In addition, there is some basis in psychological theory for extrapolating from the behavior of laypeople to that of experts. What follows is a selection of the kinds of problems that any of us may encounter when going beyond the available data, and which must be considered when weighing the usefulness of analyses estimating risks and benefits.

Sensitivity to Sample Size

Tversky and Kahneman (1971) found that even statistically sophisticated individuals have poor intuitions about the size of sample

needed to test research hypotheses adequately. In particular, they expect small samples to represent the populations from which they were drawn to a degree that can only be assumed with much larger samples. This tendency leads them to gamble their research hypotheses on underpowered small samples, to place undue confidence in early data trends, and to underestimate the role of sampling variability in causing results to deviate from expectations (preferring instead to offer causal explanations for discrepancies). For example, in a survey of standard hematology texts, Berkson et al. (1939–1940) found that the maximum allowable difference between two successive blood counts was so small that it would normally be exceeded by chance 66 to 85 percent of the time. They mused about why instructors often reported that their best students had the most trouble attaining the desired standard.

Small samples mean low statistical power, that is, a small chance of detecting phenomena that really exist. Cohen (1962) surveyed published articles in a respected psychological journal and found very low power. Even under the charitable assumption that all underlying effects were large, a quarter of the studies had less than three chances in four of showing statistically significant results. He goes on to speculate that the one way to get a low-power study published is to keep doing it again and again (perhaps making subtle variations designed to "get it right next time") until a significant result occurs. Consequently, published studies may be unrepresentative of the set of conducted studies in a way that inflates the rate of spuriously significant results (beyond that implied by the officially reported "significance level"). Page (1981) has similarly shown the low power of representative toxicological studies. In designing such studies, one inevitably must make a trade-off between avoiding false alarms (e.g., erroneously calling a chemical a carcinogen) and misses (e.g., erroneously calling a chemical a noncarcinogen). Low power increases the miss rate and decreases the false alarm rate. Hence, wayward intuitions may lead to experimental designs that represent, perhaps inadvertently, a social policy that protects chemicals more than people.

Hindsight

Experimental work has shown that in hindsight people consistently exaggerate what could have been anticipated in foresight.

They tend not only to view what has happened as having been relatively inevitable, but also to view it as having appeared relatively inevitable before it happened. People believe that others should have been able to anticipate events much better than was actually the case. They even misremember their own predictions so as to exaggerate in hindsight what they knew in foresight (Fischhoff, 1980).

The revisionist history of strategic surprises (e.g., Lanir, 1982; Wohlstetter, 1962) argues that such misperceptions have vitiated the efforts of scholars and "scalpers" attempting to understand questions like, "Who goofed at Pearl Harbor?" These expert scrutinizers were not able to disregard the knowledge that they had only as a result of knowing how things turned out. Although it is flattering to believe that we personally would not have been surprised, failing to realize the difficulty of the task that faced the individuals about whom we are speculating may leave us very exposed to future surprises.

Methodological treatises for professional historians contain numerous warnings about related tendencies. One such tendency is telescoping the rate of historical processes, exaggerating the speed with which "inevitable" changes are consummated (Fischer, 1970). Mass immunization against poliomyelitis seems like such a natural idea that careful research is needed to show that its adoption met substantial snags, taking almost a decade to complete (Lawless, 1977). A second variant of hindsight bias may be seen in Barraclough's (1972) critique of the historiography of the ideological roots of Nazism; looking back from the Third Reich, one can trace its roots to the writings of many authors from whose writings one could not have projected Nazism. A third form of hindsight bias, also called "presentism," is to imagine that the participants in a historical situation were fully aware of its eventual importance ["Dear Diary, The Hundred Years' War started today" (Fischer, 1970)].

More directly relevant to the resolution of scientific disputes, Lakatos (1970) has argued that the "critical experiment," unequivocally resolving the conflict between two theories or establishing the validity of one, is typically an artifact of inappropriate reconstruction. In fact, "the crucial experiment is seen as crucial only decades later. Theories don't just give up, a few anomalies are always allowed. Indeed, it is very difficult to defeat a research programme supported by talented and imaginative scientists" (Lakatos, 1970:157–158).

Future generations may be puzzled by the persistence of the antinuclear movement after the 1973 Arab oil embargo guaranteed the future of nuclear power, or the persistence of nuclear advocates

after Three Mile Island sealed the industry's fate—depending on how things turn out. Perhaps the best way to protect ourselves from the surprises and reprobation of the future in managing hazards is to "accept the fact of uncertainty and learn to live with it. Since no magic will provide certainty, our plans must work without it" (Wohlstetter, 1962:401).

Judging Probabilistic Processes

After seeing four successive heads in flips of a fair coin, most people expect a tails. Once diagnosed, this tendency is readily interpreted as a judgmental error. Commonly labeled the "gambler's fallacy" (Lindman and Edwards, 1961), it is one reflection of a strong psychological tendency to impose order on the results of random processes, making them appear interpretable and predictable (Kahneman and Tversky, 1972). Such illusions need not disappear with higher stakes or greater attention to detail. Feller (1968) offers one example in risk monitoring: Londoners during the Blitz devoted considerable effort to interpreting the pattern of German bombing, developing elaborate theories of where the Germans were aiming (and when to take cover). However, a careful statistical analysis revealed that the frequency distribution of bomb-hits in different sections of London was almost a perfect approximation of the Poisson (random) distribution. Dreman (1979) argues that the technical analysis of stock prices by market experts represents little more than opportunistic explication of chance fluctuations. Although such predictions generate an aura of knowing, they fail to outperform market averages.

Gilovich et al. (1985) found that, appearances to the contrary, basketball players have no more shooting streaks than one might expect from a random process generated by their overall shooting percentage. This result runs strongly counter to the conventional wisdom that players periodically have a "hot hand," attributable to specific causes like a half-time talk or dedication to an injured teammate. One of the few basketball experts to accept this result claimed that he could not act on it anyway. Fans would not forgive him if, in the closing minutes of a game, he had an inbound pass directed to a higher percentage shooter, rather than to a player with an apparent "hot hand" (even knowing that opposing players would cluster on that player, expecting the pass).

At times, even scientific enterprises seem to represent little more than sophisticated capitalization on chance. Chapman and Chapman (1969) found that clinical psychologists see patterns that they expect to find even in randomly generated data. O'Leary et al. (1974) observed that the theories of foreign affairs analysts are so complicated that any imaginable set of data can be interpreted as being consistent with them. Short of this extreme, it is generally true that, given a set of events (e.g., environmental calamities) and a sufficiently large set of possible explanatory variables (antecedent conditions), one can always devise a theory for retrospectively predicting the events to any desired level of proficiency. The price one pays for such overfitting is shrinkage, failure of the theory to work on a new sample of cases. The frequency and vehemence of warnings against such correlational overkill suggest that this bias is quite resistant to even extended professional training (Armstrong, 1975; Campbell, 1975; Crask and Parreault, 1977; Kunce et al., 1975).

Even when one is alert to such problems, it may be difficult to assess the degree to which one has capitalized on chance. For example, as a toxicologist, you are "certain" that exposure to chemical X is bad for one's health, so you compare workers who do and do not work with it in a particular plant for bladder cancer, but obtain no effect. So you try intestinal cancer, emphysema, dizziness, and so on, until you finally get a significant difference in skin cancer. Is that difference meaningful? Of course, the way to test these explanations or theories is by replication on new samples. That step, unfortunately, is seldom taken and is often not possible for technical or ethical reasons (Tukey, 1977).

A further unintuitive property of probabilistic events is regression to the mean, the tendency for extreme observations to be followed by less extreme ones. One depressing failure by experts to appreciate this fact is seen in Campbell and Erlebacher's (1970) article, "How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful" (because upon retest, the performance of the better students seems to have deteriorated). Similarly unfair tests may be created when one asks only if environmental management programs have, say, weakened strong industries or reduced productivity in the healthiest sectors of the economy.

Judging the Quality of Evidence

Since cognitive and evidential limits prevent scientists from providing all the answers, it is important to have an appraisal of how much they do know. It is not enough to claim that "these are the ranking experts in the field," for there are some fields in which the most knowledgeable individuals understand a relatively small portion of all there is to be known.

Weather forecasters offer some reason for encouragement (Murphy and Brown, 1983; Murphy and Winkler, 1984). There is at least some measurable precipitation on about 70 percent of the occasions for which they say there is a 70 percent chance of rain. The conditions under which forecasters work and train suggest the following prerequisites for good performance in probabilistic judgment:

- great amounts of practice;
- the availability of statistical data offering historical precipitation base rates (indeed, forecasters might be fairly well calibrated if they ignored the murmurings of their intuitions and always responded with the base rate);
- computer-generated predictions for each situation;
- a readily verifiable criterion event (measurable precipitation), offering clear feedback; and
- explicit admission of the imprecision of the trade and the need for training.

In experimental work, it has been found that large amounts of clearly characterized, accurate, and personalized feedback can improve the probability assessments of laypeople (e.g., Lichtenstein and Fischhoff, 1980).

Training professionals to assess and express their uncertainty is, however, a rarity. Indeed, the role of judgment is often acknowledged only obliquely. For example, civil engineers do not routinely assess the probability of failure for completed dams, even though approximately one dam in 300 collapses when first filled (U.S. Committee on Government Operations, 1978). The "Rasmussen" Reactor Safety Study (U.S. Nuclear Regulatory Commission, 1975) was an important step toward formalizing the role of risk in technological systems, although a subsequent review was needed to clarify the extent to which these estimates were but the product of fallible, educated judgment (U.S. Nuclear Regulatory Commission, 1978).

Ultimately, the quality of experts' assessments is a matter of judgment. Since expertise is so narrowly distributed, assessors are typically called upon to judge the quality of their own judgments. Unfortunately, an extensive body of research suggests that people are overconfident when making such assessments (Lichtenstein et al., 1982). A major source of such overconfidence seems to be failure to appreciate the nature and tenuousness of the assumptions on which judgments are based. To illustrate with a trivial example, when asked "To which country are potatoes native? (a) Ireland (b) Peru?", many people are very confident that answer (a) is true. The Irish potato and potato blight are familiar to most people; however, that is no guarantee of origin. Indeed, the fact that potatoes were not indigenous to Ireland may have increased their susceptibility to blight there.

Experts may be as prone to overconfidence as laypeople (in cases in which they, too, are pressed to evaluate judgments made regarding topics about which their knowledge is limited). For example, when several internationally known geotechnical experts were asked to predict the height of fill at which an embankment would fail and to give confidence intervals for their estimates, without exception, the true values fell outside the confidence intervals (Hynes and Vanmarcke, 1976), a result akin to that observed with other tasks and respondent populations (Lichtenstein et al., 1982). One of the intellectual challenges facing engineering is to systematize the role of judgment, both to improve its quality and to inform those who must rely on it in their decision making.

This basic pattern of results has proved so robust that it is hard to acquire much insight into the psychological processes producing it (Lichtenstein et al., 1982). One of the few effective manipulations is to force subjects to explain why their chosen answers might be wrong (Koriat et al., 1980). That simple instruction seems to prompt recall of contrary reasons that would not normally come to mind given people's natural thought processes, which seem to focus on retrieving reasons that support chosen answers. A second seemingly effective manipulation, mentioned earlier, is to train people intensively with personalized feedback that shows them how well they are calibrated.

Figures II.9 and II.10 show one sign of the limits that exist on the capacity of expertise and experience to improve judgment—in the absence of the conditions for learning enjoyed, say, by weather forecasters. Particle physicists' estimates of the value of several physical constants are bracketed by what might be called confidence intervals, showing the range of likely values within which the true

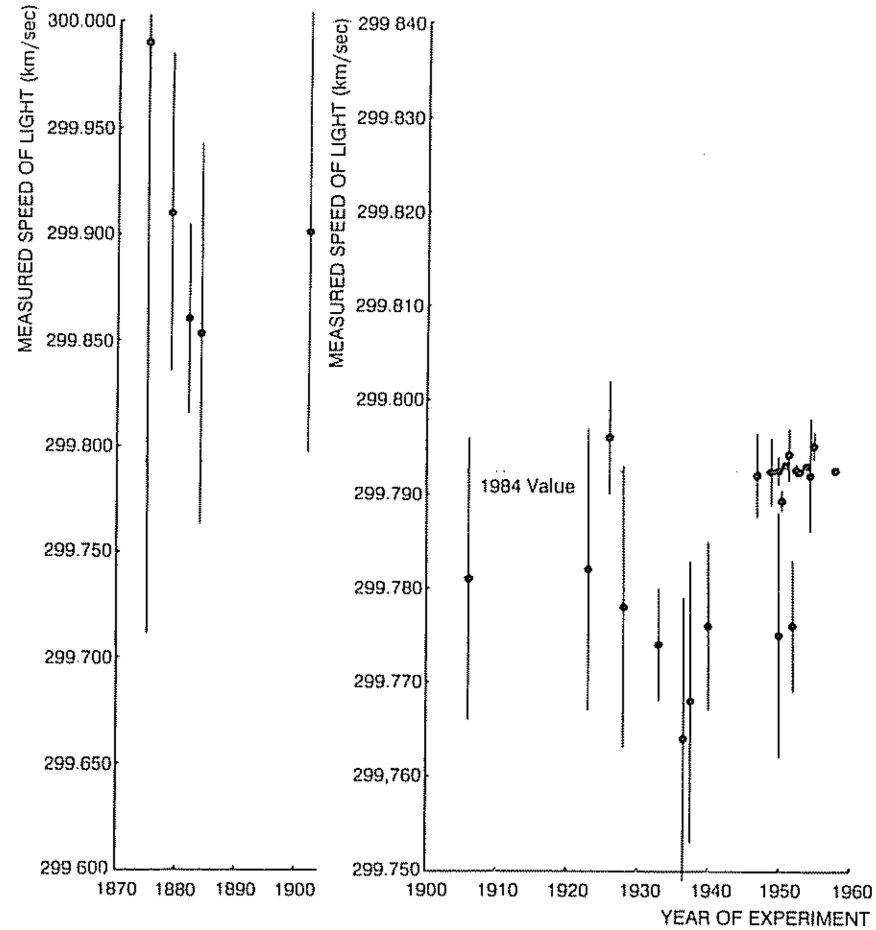

FIGURE II.9 Calibration of confidence in estimates of physical constants. SOURCE: Henrion and Fischhoff, 1986. Copyright © 1986 by the American Association of Physics Teachers.

value should fall, once it is known. Narrower intervals indicate greater confidence. These intervals have shrunk over time, as physicists' knowledge has increased. However, at most points, they seem to have been too narrow. Otherwise, the new best estimates would not have fallen so frequently outside the range of what previously seemed plausible. In an absolute sense, the level of knowledge represented here is extremely high and the successive best estimates lie extremely close to one another. However, the confidence intervals define what constitute surprises in terms of current physical theory. Unless the

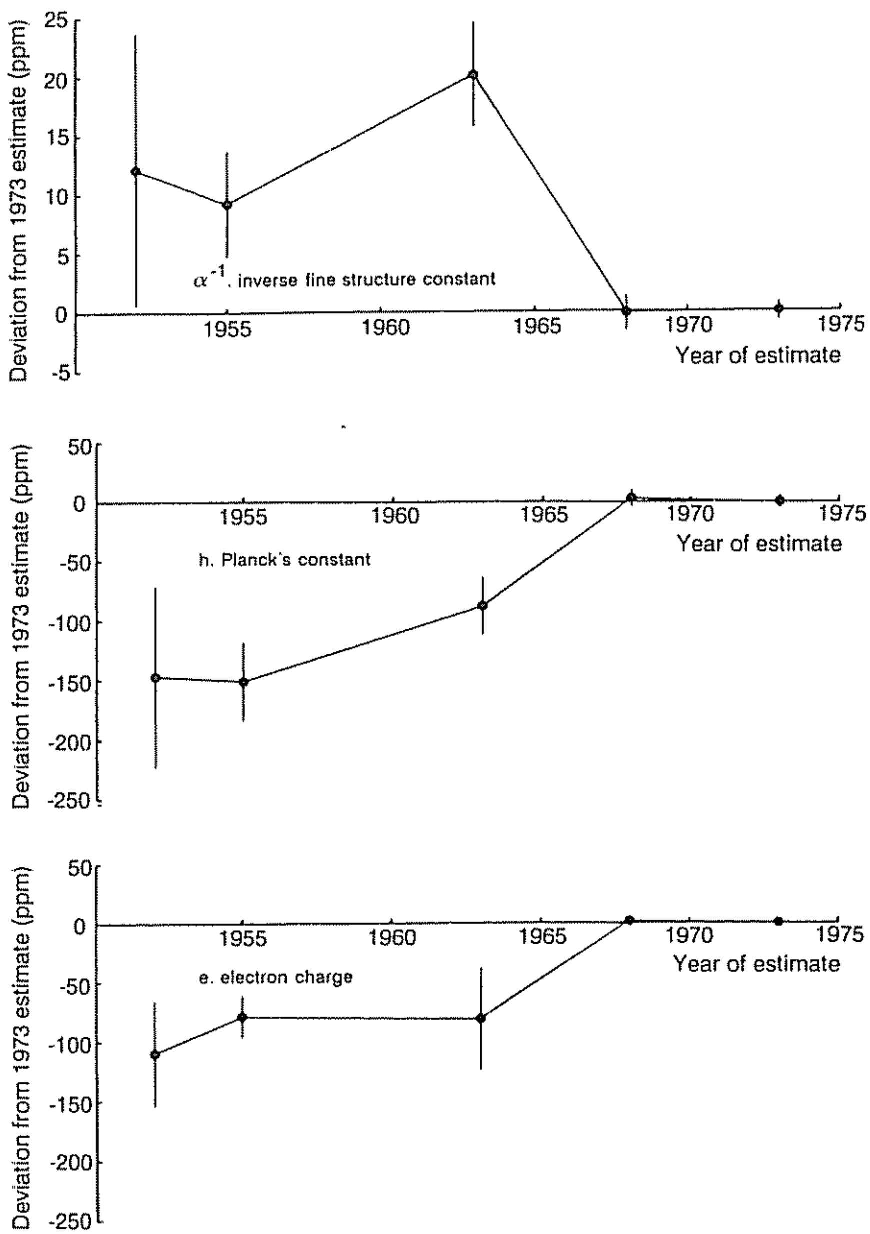

FIGURE II.10 Recommended values for fundamental constants, 1952 through 1973. SOURCE: Henrion and Fischhoff, 1986. Copyright © 1986 by the American Association of Physics Teachers.

possibility of overconfident judgment is considered, values falling outside the intervals suggest a weakness in theory.

SUMMARY

The science of risk provides a critical anchor for risk controversies. There is no substitute for that science. However, it is typically an imperfect guide. It can mislead if one violates any of a wide variety of intricate methodological requirements—including the need to use judgment judiciously (and to understand its limitations). The general nature of these assumptions was illustrated with examples drawn from the science of understanding human behavior. Sections IV through VI deal with the human anchors for risk controversies: the nature of their political tensions, the strategies that risk communicators can take in them, and psychological barriers to risk communication. The next section (III) deals with the interface between science and behavior, specifically ways in which science shapes and is shaped by the political process.

III
SCIENCE AND POLICY

SEPARATING FACTS AND VALUES

The first recommendation of the National Research Council's Committee on the Institutional Means for Assessment of Risks to Public Health (National Research Council, 1983b:7) was that:

> regulatory agencies take steps to establish and maintain a clear conceptual distinction between assessment of risks and considerations of risk management alternatives; that is, the scientific findings and policy judgments embodied in risk assessments should be explicitly distinguished from the political, economic, and technical considerations that influence the design and choice of regulatory strategies.

The principle of separating science and politics seems to be a cornerstone of professional risk management. Many of the antagonisms surrounding risk management seem due to the blurring of this distinction, resulting in situations in which science is rejected because it is seen as tainted by politics. As Hammond and Adelman (1976), Mazur et al. (1979), and others have argued, this distinction can help clear the air in debates about risk, which might otherwise fill up with half-truths, loaded language, and character assassinations. Even technical experts may fall prey to partisanship as they advance views on political topics beyond their fields of expertise, downplay facts they believe will worry the public, or make statements that cannot be verified.

Although a careful delineation between values and facts can help prevent values from hiding in facts' clothing, it cannot assure that a complete separation will ever be possible (Bazelon, 1979; Callen, 1976). The "facts" of a matter are only those deemed relevant to a particular problem, whose definition forecloses some action options and effectively prejudges others. Deciding what the problem is goes a long way to determining what the answer will be. Hence, the "objectivity" of the facts is always conditioned on the assumption that they are addressing the "right" problem, where "right" is defined in terms of society's best interest, not the interest of a particular party. The remainder of this section examines how our values determine what facts we produce and use, and how our facts shape our values.

Values Shape Facts

Without information, it may be hard to arouse concern about an issue, to allay fears, or to justify an action. But information is usually created only if someone has a use for it. That use may be pecuniary, scientific, or political. Thus, we may know something only if someone in a position to decide feels that it is worth knowing. Doern (1978) proposed that lack of interest in the fate of workers was responsible for the lack of research on the risks of uranium mining; Neyman (1979) wondered whether the special concern with radiation hazards had restricted the study of chemical carcinogens; Commoner (1979) accused oil interests of preventing the research that could establish solar power as an energy option. In some situations, knowledge is so specialized that all relevant experts may be in the employ of a technology's promoters, leaving no one competent to discover troublesome facts (Gamble, 1978). Conversely, if one looks hard enough for, say, adverse effects of a chemical, chance alone will produce an occasional positive finding. Although such spurious results are likely to vanish when studies are replicated, replications are the exception rather than the rule in many areas. Moreover, the concern raised by a faulty study may not be as readily erased from people's consciousness as from the scientific literature (Holden, 1980; Kolata, 1980; Peto, 1980). A shadow of doubt is hard to remove.

Legal requirements are an expression of society's values that may strongly affect its view of reality. Highway-safety legislation affects accident reports in ways that are independent of its effects on accident rates (V.L. Wilson, 1980). Crime-prevention programs may have similar effects, inflating the perceived problem by encouraging victims to report crimes (National Research Council, 1976). Although it is not always exploited for research purposes, an enormous legacy of medical tests has been created by the defensive medicine engendered by fear of malpractice. Legal concerns may also lead to the suppression of information, as doctors destroy "old" records that implicate them in the administration of diethylstilbestrol (DES) to pregnant women in the 1950s, employers fail to keep "unnecessary" records on occupational hazards, or innovators protect proprietary information (Lave, 1978; Pearce, 1979; Schneiderman, 1980).

Whereas individual scientists create data, it is the community of scientists and other interpreters who create facts by integrating data (Levine, 1974). Survival in this adversarial context is determined in part by what is right (i.e., truth) and in part by the staying power of those who collect particular data or want to believe in them. Scrutiny from both sides in a dispute is a valuable safeguard, likely to improve the quality of the analysis. Each side tries to eliminate erroneous material prejudicial to its position. If only one side scrutinizes, the resulting analyses will be unbalanced. Because staying with a problem requires resources, the winners in the marketplace of ideas may tend to be the winners in the political and economic marketplace.

Facts Shape Values

Values are acquired by rote (e.g., in Sunday school), by imitation, and by experience (Rokeach, 1973). The world we observe tells us what issues are worth worrying about, what desires are capable of fruition, and who we are in relation to our fellows. Insofar as that world is revealed to us through the prism of science, the facts it creates help shape our world outlook (R.P. Applebaum, 1977; Henshel, 1975; Markovic, 1970; Shroyer, 1970). The content of science's facts can make us feel like hedonistic consumers wrestling with our fellows, like passive servants of society's institutions, like beings at war with or at one with nature. The quantity of science's facts (and the coherence of their explication) may lower our self-esteem and enhance that of technical elites. The topics of science's inquiries may tell us that the important issues of life concern the mastery of others and of nature, or the building of humane relationships. Some argue that science can "anaesthetize moral feeling" (Tribe, 1972) by enticing us to think about the unthinkable. For example, setting an explicit value on human life in order to guide policy decisions may erode our social contract, even though we set such values implicitly by whatever decisions we make.

Even flawed science may shape our values. According to Wortman (1975), Westinghouse's poor evaluation of the Head Start program in the mid-1960s had a major corrosive effect on faith in social programs and liberal ideals. Weaver (1979) argued that whatever technical problems may be found with Inhaber's (1979) comparison of the risks of different energy sources, he succeeded in creating a new perspective that was deleterious to the opponents of nuclear power. As mentioned earlier, incorrect intuitions regarding the statistical power of statistical designs can lead to research that implicitly values chemicals more than people (Page, 1978, 1981). In designing such studies, one must make a trade-off between avoiding either false alarms (e.g., erroneously calling a chemical a carcinogen) or misses (e.g., not identifying a carcinogen as such). The decision to study

many chemicals with relatively small samples both increases the miss rate and decreases the false-alarm rate. The value bias of such studies is compounded when scientific caution also becomes regulatory caution.

Where science concerns real-world objects, then the selection and characterization of those objects inevitably express attitudes toward them. Those attitudes may come from the risk managers who commission scientific studies, or they may come from the scientists who conduct them. In either case, the deepest link between science and politics may be in basic issues of definition. The next section discusses some of the subtle ways in which science can preempt or be captured by the policymaking process in its treatment of two basic concepts of risk management: risk and benefit.

MEASURING RISK

Which Hazards Are Being Considered?

The decision to decide whether a technology's risks are acceptable implies that, in the opinion of someone who matters, it may be too dangerous. Such issue identification is itself an action with potentially important consequences. Putting a technology on the decision-making agenda can materially change its fate by attracting attention to it and encouraging the neglect of other hazards. For example, concern about carbon-dioxide-induced climatic change (Schneider and Mesirow, 1976) changes the status of fossil fuels vis-à-vis nuclear power.

After an issue has been identified, the hazard in question must still be defined. Breadth of definition is particularly important. Are military and nonmilitary nuclear wastes to be lumped together in one broad category, or do they constitute separate hazards? Did the collision of two jumbo jets at Tenerife in the Canary Islands represent a unique miscommunication or a large class of pilot–controller impediments? Do all uses of asbestos make up a single industry or are brake linings, insulation, and so forth to be treated separately? Do hazardous wastes include residential sewage or only industrial solids (*Chemical and Engineering News*, 1980)? Grouping may convert a set of minor hazards into a major societal problem, or vice versa. Lead in the environment may seem worth worrying about, but lead solder in tuna fish cans may not. In recent years, isolated cases of child abuse have been aggregated in such a way that a persistent problem with a relatively stable rate of occurrence now appears as an epidemic demanding action.

Often the breadth of a hazard category becomes apparent only after the decision has been made and its implications experienced in practice. Some categories are broadened, for example, when precedent-setting decisions are applied to previously unrelated hazards. Other categories are narrowed over time as vested interests gain exceptions to the rules applying to the category in which their technology once belonged (Barber, 1979). In either case, different decisions might have been made had the hazard category been better defined in advance.

Definition of Risk

Managing technological risks has become a major topic in scientific, industrial, and public policy. It has spurred the development of some industries and prompted the demise of others. It has expanded the powers of some agencies and overwhelmed the capacity of others. It has enhanced the growth of some disciplines and changed the paths of others. It has generated political campaigns and countercampaigns. The focal ingredient in all this has been concern over risk. Yet, the meaning of "risk" has always been fraught with confusion and controversy. Some of this conflict has been overt, as when a professional body argues about the proper measure of pollution or reliability for incorporation in a health or safety standard. More often, though, the controversy is unrecognized; the term risk is used in a particular way without extensive deliberations regarding the implications of alternative uses. Typically, that particular way follows custom in the scientific discipline initially concerned with the risk.

However, the definition of risk, like that of any other key term in policy issues, is inherently controversial. The choice of definition can affect the outcome of policy debates, the allocation of resources among safety measures, and the distribution of political power in society.

Dimensionality of Risk

The risks of a technology are seldom its only consequences. No one would produce it if it did not generate some benefits for someone. No one could produce it without incurring some costs. The difference between these benefits and nonrisk costs could be called

the technology's net benefit. In addition, risk itself is seldom just a single consequence. A technology may be capable of causing fatalities in several ways (e.g., by explosions and chronic toxicity), as well as inducing various forms of morbidity. It can affect plants and animals as well as humans. An analysis of risk needs to specify which of these dimensions will be included. In general, definitions based on a single dimension will favor technologies that do their harm in a variety of ways (as opposed to those that create a lot of one kind of problem). Although it represents particular values (and leads to decisions consonant with those values), the specification of dimensionality (like any other specification) is often the inadvertent product of convention or other forces, such as jurisdictional boundaries (Fischhoff, 1984).

Summary Statistics

For each dimension selected as relevant, some quantitative summary is needed for expressing how much of that kind of risk is created by a technology. The controversial aspects of that choice can be seen by comparing the practices of different scientists. For some, the unit of choice is the annual death toll (e.g., Zentner, 1979); for others, deaths per person exposed or per hour of exposure (e.g., Starr, 1969); for others, it is the loss of life expectancy (e.g., Cohen and Lee, 1979; Reissland and Harries, 1979); for still others, lost working days (e.g., Inhaber, 1979). Crouch and Wilson (1982) have shown how the choice of unit can affect the relative riskiness of technologies. For example, today's coal mines are much less risky than those of 30 years ago in terms of accidental deaths per ton of coal, but marginally riskier in terms of accidental deaths per employee. The difference between measures is explained by increased productivity. The choice among measures is a policy question, with Crouch and Wilson suggesting that:

> From a national point of view, given that a certain amount of coal has to be obtained, deaths per million tons of coal is the more appropriate measure of risk, whereas from a labor leader's point of view, deaths per thousand persons employed may be more relevant (1982:13).

Other value questions may be seen in the units themselves. For example, loss of life expectancy places a premium on early deaths that is absent from measures treating all deaths equally; using it means ascribing particular worth to the lives of young people. Just counting fatalities expresses indifference to whether they come immediately after mishaps or following a substantial latency period (during which it may not be clear who will die). Whatever types of individuals are included in a category, they are treated as equals; the categories may include beneficiaries and nonbeneficiaries of a technology (reflecting an attitude toward that kind of equity), workers and members of the general public (reflecting an attitude toward that kind of voluntariness), or participants and nonparticipants in setting policy for the technology (reflecting an attitude toward that kind of voluntariness). Using the average of past casualties or the expectation of future fatalities means ignoring the distribution of risk over time; it treats technologies taking a steady annual toll in the same way as those that are typically benign, except for the rare catastrophic accident. When averages are inadequate, a case might be made for using one of the higher moments of the distribution of casualties over time or for incorporating a measure of the uncertainty surrounding estimates (Fischhoff, 1984).

Bounding the Technology

Willingness to count delayed fatalities means that a technology's effects are not being bounded in time (as they are, for example, in some legal proceedings that consider the time that passes between cause, effect, discovery, and reporting). Other bounds need to be set also, either implicitly or explicitly. One is the proportion of the fuel and materials cycles to be considered: To what extent should the risks be restricted to those people who enjoy the direct benefits of a technology or extended to cover those involved in the full range of activities necessary if those benefits are to be obtained? Crouch and Wilson (1982) offer an insightful discussion of some of these issues in the context of imported steel; the U.S. Nuclear Regulatory Commission (1983) has adopted a restrictive definition in setting safety goals for nuclear power (Fischhoff, 1983); much of the acrimony in the debates over the risks of competing energy technologies concerned treatment of the risks of back-up energy sources (Herbert et al., 1979; Inhaber, 1979). A second recurrent bounding problem is how far to go in considering higher-order consequences (i.e., when coping with one risk exposes people to another). As shown in Figure II.1, hazards begin with the human need the technology is designed to satisfy, and develop over time. One can look at the whole process or only at its conclusion. The more narrowly a hazard's moment in time is

defined, the fewer the options that can be considered for managing its risks. A third issue of limits is how to treat a technology's partial contribution to consequences, for example, when it renders people susceptible to other problems or when it accentuates other effects through synergistic processes.

Concern

Events that threaten people's health and safety exact a toll even if they never happen. Concerns over accidents, illness, and unemployment occupy people even when they and their loved ones experience long, robust, and salaried lives. Although associated with risks, these consequences are virtual certainties. All those who know about them will respond to them in some way. In some cases, that response benefits the respondent, even if its source is an aversive event. For example, financial worries may prompt people to expand their personal skills or create socially useful innovations. Nonetheless, their resources have been diverted from other, perhaps preferred pursuits. Moreover, the accompanying stress can contribute to a variety of negative health effects, particularly when it is hard to control the threat (Elliot and Eisdorfer, 1982). Stress not only precipitates problems of its own, but can complicate other problems and divert the psychological resources needed to cope with them. Thus, concern about a risk may hasten the end of a marriage by giving the couple one more thing to fight about and that much less energy to look for solutions.

Hazardous technologies can evoke such concern even when they are functioning perfectly. Some of the response may be focused and purposeful, such as attempts to reduce the risk through personal and collective action. However, even that effort should be considered a cost of the technology because that time and energy might have been invested in something else (e.g., leisure, financial planning, improving professional skills) were it not for the technology. When many people are exposed to the risk (or are concerned about the exposure of their fellows), then the costs may be extensive. Concern may have even greater impact than the actual health and safety effects of the technology. Ironically, because the signs of stress are diffuse (e.g., a few more divorces, somewhat aggravated cardiovascular problems), it is quite possible for the size of the effects to be both intolerably large (considering the benefits) and undetectable (by current techniques).

Including concern among the consequences of a risky technology immediately raises two additional controversial issues. One centers on what constitutes an appropriate level of concern. It could be argued that concern should be proportionate to physical risk. There are, however, a variety of reasons why citizens might reasonably be concerned most about hazards that they themselves acknowledge to be relatively small (e.g., they feel that an important precedent is being set, that things will get worse if not checked, or that the chances for effective action are great) (see Section IV). The second issue is whether to hold a technology responsible for the concern evoked by people's perceptions of its risks or for the concern that would be evoked were people to share the best available technical knowledge. It is the former that determines actual concern; however, using it would mean penalizing some technologies for evoking unjustified concerns and rewarding others for having escaped the public eye.

MEASURING BENEFITS

Although the term risk management is commonly used for dealing with potentially hazardous technologies, few risk policies are concerned entirely with risk. Technologies would not be tolerated if they did not bring some benefit. Residual risk would not be tolerated if the benefits of additional reduction did not seem unduly expensive (to whoever is making the decision). As a result, some assessment of benefits is a part of all risk decisions, whether undertaken by institutions or by individuals. Faith in quantification makes formal cost–benefit analysis a part of many governmental decisions in the United States (Bentkover et al., 1985). However, a variety of procedures are possible, each with its own behavioral and ethical assumptions.

Definition of Benefit

Benefit assessment begins with a series of decisions that bound the analysis and specify its key terms. Together, these decisions provide an operational definition of what "benefit" means. Although they may seem technical and are often treated in passing, these decisions are the heart of an analysis. They express a social philosophy, elaborating what society holds to be important in a particular context. The ensuing analysis is "merely" an exercise in determining how well different policy options realize this philosophy. If the philosophy has not been interpreted, stated, and implemented appropriately, then the analysis becomes an exercise in futility.

The details of this definitional process in some ways parallel that for defining risk. Policymakers commission benefit assessments to help them make decisions; that is, to help them choose among alternative courses of action (including, typically, inaction). To make those decisions, they must (1) identify the policy alternatives (or options) that could be adopted; (2) circumscribe the set of policy-relevant consequences that these alternatives could create; (3) estimate the magnitude of each alternative's consequences were it adopted; (4) evaluate the benefits (and costs) that affected individuals would derive from these consequences; and (5) aggregate benefits across individuals. Defining the policymaking question is a precondition for commissioning any benefit assessment meant to serve it. For example, one cannot calculate the consequences of one particular policy without knowing the alternative policies that might come in its stead were it not adopted (and whose benefits would be foregone if it was). One cannot begin to assess and tally benefits without knowing which consequences and individuals fall within the agency's jurisdiction. Figure III.1 provides a summary of these definitional issues. Fischhoff and Cox (1985) discuss them in greater detail.

Once it has been determined what evaluations to seek, a method must be found for doing the seeking. There are two natural places to look for guidance regarding the evaluation of benefits: what people say and what people do. Methods relying on the former consider expressed preferences; methods relying on the latter consider revealed preferences. Each makes certain ethical and empirical assumptions regarding the nature of individual and societal behavior, the validity of which determines their applicability to particular situations (Driver et al., 1988).

Expressed Preferences

The most straightforward way to find out what people value, regarding safety or anything else, is to ask them. The asking can be done at the level of overall assessments (e.g., "Do you favor . . . ?"), statements of principle (e.g., "Should our society be risk averse regarding . . . ?"), or detailed trade-offs (e.g., "How much of a monetary sacrifice would you make in order to ensure . . . ?"). The vehicle for collecting these values could be public opinion polls (Conn, 1983), comments solicited at public hearings (Mazur, 1973; Nelkin, 1984), or detailed interviews conducted by decision analysts or counselors (Janis, 1982; Keeney, 1980). The advantages of these

IDENTIFYING THE SET OF POLICY OPTIONS
- Specifying details of each option
- Determining the range of variation
- Assessing the uncertainty surrounding implementation
- Anticipating the stability of the situation following inaction
- Determining the legitimacy of creating new options arising during the analysis

IDENTIFYING THE SET OF RELEVANT CONSEQUENCES
- Choosing consequences
  - Scientific, legal, political, ethical grounds
  - Public and private goods
- Specifying consequences
  - Bounding in space
  - Bounding in time
  - Including higher-order consequences
  - Including associated concern

ESTIMATING THE MAGNITUDES OF CONSEQUENCES
- Assessing the uncertainty around estimates
- Determining the risk assessor's attitude toward uncertainty
- Identifying deliberate bias in estimates
- Discerning the presuppositions in terms

EVALUATING BENEFITS FOR INDIVIDUALS
- Defining individuals
- Determining initial entitlements (willingness to pay versus willingness to accept)
- Identifying ultimate arbiter of benefit

AGGREGATING NET BENEFITS ACROSS INDIVIDUALS
- Looking for dominating alternatives (Pareto optimality)
- Exploring utilitarian solutions (potential Pareto improvements)
- Using group utility functions
- Resolving distributional inequities

FIGURE III.1 Steps in problem definition. SOURCE: Fischhoff and Cox, 1985.

procedures are that they are current (in the sense of capturing today's values), sensitive (in the sense of theoretically allowing people to say whatever they want), specifiable (in the sense of allowing one to ask the precise questions that interest policymakers), direct (in the sense of looking at the preferences themselves and not how they reveal themselves in application to some specific decision problem), superficially simple (in the sense that you just ask people questions), politically appealing (in the sense that they let "the people" speak), and instructive (in the sense that they force people to think in a focused manner about topics that they might otherwise ignore).

As discussed in Section II, however, a number of difficult conditions must be met if expressed preference procedures are to fulfill their promise. One is that the question asked must be the precise one needed for policymaking (e.g., "How much should you be paid in order to incur a 10 percent increase in your annual probability of an injury sufficiently severe to require at least one day of hospitalization, but not involving permanent disability?"), rather than an ill-defined one, such as "do you favor better roads?" or "is your job too risky?" (In response, a thoughtful interviewee might ask, "What alternatives should I be considering? Am I allowed to consider who pays for improvements?") One response to the threat of ambiguity is to lay out all details of the evaluation question to respondents (Fischhoff and Furby, 1988). A threat to this solution is that the full specification will be so complex and unfamiliar as to pose an overwhelming inferential task. To avoid the incompletely considered, and potentially labile, responses that might arise, one must either adjust the questions to the respondents or the respondents to the questions. The former requires an empirically grounded understanding of what issues people have considered and how they have thought about them. This understanding allows one to focus the interview on the areas in which people have articulated beliefs, to provide needed elaborations, and to avoid repeating details that correspond to respondents' default assumptions (and could, therefore, go without saying).

If the gap between policymakers' questions and respondents' answers is too great to be bridged in a standard interviewing session, then it may be necessary either to simplify the questions or to complicate the session. A structured form of simplification is offered by techniques, such as multi-attribute utility theory, which decompose complex questions into more manageable components, each of which considers a subsidiary evaluation issue (Keeney and Raiffa, 1976). The structuring of these questions allows their recomposition into overall evaluations, which are interpreted as representing the summary judgments that respondents would have produced if they had unlimited mental computational capacity. The price paid for this potential simplification is the need to answer large numbers of simple, formal, and precise questions.

Where it becomes impossible to bring the question "down" to the level of the respondent, there still may be some opportunity to bring the respondent "up" to the level of the question. Ways of enabling respondents to realize their latent capability for thinking meaningfully about questions include talking with them about the issues, including them in focused group discussions, suggesting alternative perspectives (for their consideration), and giving them time to ruminate over their answers.

Revealed Preferences

The alternative to words is action. This collection of techniques assumes that people's overt actions can be interpreted to reveal the preferences that motivated them. The great attraction of such procedures is that they are based on real acts, whose consequences are presumably weightier than those of even the most intelligently conducted interview. They focus on possibilities, rather than just desires.

By concentrating on current, real decisions, these procedures are also strongly anchored in the status quo. It is today's work, with today's constraints, that conditions the behavior observed. If today's society inhibits people's ability to act in ways that express their fundamental values, then revealed preference procedures lose their credibility (whereas expressed preferences, at least in principle, allow people to raise themselves above today's reality). Thus, if one feels that advertising, or regulation, or monopoly pressures have distorted contemporary evaluations of some products or consequences, then revealing those values does not yield a guide to true worth. Relying on those values for policymaking would mean enshrining today's imperfections (and inequities) in tomorrow's world.

The commitment to observing actual behavior also makes these procedures particularly vulnerable to deviations from optimality. A much smaller set of inferences separates people's true values from their expressed preferences than from their overt behavior. On the one hand, this means that people must complete an even more complex series of inferences in order to do what they want than to say what they want. On the other hand, investigators must make even more assumptions in order to infer underlying values from what they observe. Thus, for example, it is difficult enough to determine how much compensation one would demand to accept an additional injury risk of magnitude X in one's job. Implementing that policy in an actual decision also requires that suitable options be available and that their consequences be accurately perceived. If those conditions of informed consent are not met, then the interpretation of pay–danger relationships may be quite tenuous. Workers may be coercing their employer into compensating them for imagined risks; or, they may be coerced into accepting minimal compensation by an employer cognizant of a depressed job market.

The most common kind of revealed preference analysis is also the most common kind of economic analysis: interpreting marketplace prices as indicating the true value of goods. If the goods whose values are of interest (e.g., health risks) are not traded directly, then a value may be inferred by conceptualizing the goods that are traded (e.g., jobs) as representing a bundle of consequences (e.g., risks, wages, status). Analytic techniques may then be used to discern the price that markets assign to each consequence individually, by looking at its role in determining the price paid for various goods that include it.

These regression-based procedures rest on a well-developed theoretical foundation describing why (under conditions of a free market, optimal decision making, and informed consent) prices should reveal the values that people ascribe to things (Bentkover et al., 1985). The same general thought has been applied heuristically in various schemes designed to discern the values revealed in decisions (ostensibly) taken by society as a whole or by individuals under less constrained conditions. These analyses include attempts to see what benefits society demands for tolerating the risks of different technologies (Starr, 1969), what risks people seem to accept in their everyday lives (B. Cohen and Lee, 1979; R. Wilson, 1979), and what levels of technological risk escape further regulation (Fischhoff, 1983; U.S. Nuclear Regulatory Commission, 1982). These attempts are typically quite ad hoc, with no detailed methodology specifying how they should be conducted. The implicit underlying theory assumes, in effect, that whatever is, is right and that present arrangements are an appropriate basis for future policies. Thus, these procedures can guide future decisions only if one believes that society as a whole currently gets what it wants, even with regard to regulated industries, unregulated semimonopolies, and poorly understood new technologies. Extracting useful information from them requires a very detailed assessment of the procedures that they use, the existing reality that they endorse, and the kinds of behavior that they study.

Ascertaining the validity of the theory underlying approaches to measuring "benefit" that assume optimality has often proven difficult, for what can best be described as philosophical reasons. Some investigators find it implausible that people do anything other than optimize their own best interest when making decisions, maintaining that society would not be functioning so well were it not for this ability. These investigators see their role as discerning what people are trying to optimize (i.e., what values they ascribe to various consequences).

The contrary position argues that this belief in optimality is tautological, in that one can always find something that people could be construed as trying to optimize. Looking at how decisions are actually made shows that they are threatened by all the problems that can afflict expressed preferences. Thus, for example, consumers may make suboptimal choices because a good is marketed in a way that evokes only a portion of their values, or because they unwittingly exaggerate their ability to control its risks (Svenson, 1981; Weinstein, 1980a).

Because of the philosophical differences between these positions, relatively little is known about the general sensitivity of conclusions drawn from analyses that assume optimality to deviations from optimality. The consumer of such analyses is left to discern how far conditions deviate from optimal decision making by informed individuals in an unconstrained marketplace and, then, how far those deviations threaten the conclusions of the analyses.

SUMMARY

Science is a product of society; as such, it reflects the values of its creators. That reflection may be deliberate, as when young people decide how to dedicate their lives and research institutes decide how to stay solvent. Or, it may be unconscious, as scientists routinely apply value-laden procedures and definitions just because that was what they learned to do in school. Conversely, society is partly a product of science. That influence may be direct, as when science shapes the conditions under which people live (e.g., how prosperous they are, what industries confront them). Or it may be indirect, as when science defines our relationship with nature or raises specific fears. Understanding these interdependencies is essential to, on the one hand, discerning the objective content versus inherently subjective science and, on the other hand, directing science to serve socially desired ends. An understanding of these relationships is also necessary to appropriately interpret the conflicts between lay and expert opinions that constitute the visible core of many risk controversies. The diagnoses of these conflicts are discussed in Section IV.

IV
THE NATURE OF THE CONTROVERSY

A public opinion survey (Harris, 1980) reported the following three results:

1. Among four "leadership groups" (top corporate executives, investors and lenders, congressional representatives, and federal regulators), 94 to 98 percent of all respondents agreed with the statement "even in areas in which the actual level of risk may have decreased in the past 20 years, our society is significantly more aware of risk."
2. Between 87 and 91 percent of those four leadership groups felt that "the mood of the country regarding risk" will have a substantial or moderate impact "on investment decisions—that is, the allocation of capital in our society in the decade ahead." (The remainder believed that it would have a minimal impact, no impact at all, or were not sure.)
3. No such consensus was found, however, when these groups were asked about the appropriateness of this concern about risk. A majority of the top corporate executives and a plurality of lenders believed that "American society is overly sensitive to risk," whereas a large majority of congressional representatives and federal regulators believed that "we are becoming more aware of risk and taking realistic precautions." A sample of the public endorsed the latter statement over the former by 78 to 15 percent.

In summary, there is great agreement that risk decisions will have a major role in shaping our society's future and that those decisions will, in turn, be shaped by public perceptions of risk. There is, however, much disagreement about the appropriateness of those perceptions. Some believe the public to be wise; others do not. These contrary beliefs imply rather different roles for public involvement in risk management. As a result, the way in which this disagreement is resolved will affect not only the fate of particular technologies, but also the fate of our society and its social organization.

To that end, various investigators have been studying how and how well people think about risks. Although the results of that research are not definitive as yet, they do clearly indicate that a careful diagnosis is needed whenever the public and the experts appear to disagree. It is seldom adequate to attribute all such discrepancies to public misperceptions of the science involved. From a factual perspective, that assumption is often wrong; from a societal perspective, it is generally corrosive by encouraging disrespect among the parties involved. When the available research data do not allow one to make a confident alternative diagnosis, a sounder assumption is that there is some method in the other party's apparent madness. This section offers some ways to find that method. Specifically, it offers six reasons why disagreements between the public and the experts need not be interpreted merely as clashes between actual and perceived risks.

THE DISTINCTION BETWEEN "ACTUAL" AND "PERCEIVED" RISKS IS MISCONCEIVED

Although there are actual risks, nobody knows what they are. All that anyone does know about risks can be classified as perceptions. Those assertions that are typically called actual risks (or facts or objective information) inevitably contain some element of judgment on the part of the scientists who produce them. In this light, what is commonly called the conflict between actual and perceived risk is better thought of as the conflict between two sets of risk perceptions: those of ranking scientists performing within their field of expertise and those of anybody else. The element of judgment is most minimal when all the experts do is to assess the competence of a particular study conducted within an established paradigm. It grows with the degree to which experts must integrate results from diverse studies or extrapolate from a domain in which results are readily obtainable to another in which they are really needed (e.g., from animal studies to human effects). Judgment becomes all when there are no (credible) available data, yet a policy decision requires some assessment of a particular fact. Section II discusses at length the trustworthiness of such judgments.

The expert opinions that make up the scientific literature aspire to be objective in two senses, neither of which can ever be achieved absolutely and neither of which is the exclusive province of technical experts. One meaning of objectivity is reproducibility: one expert should be able to repeat another's study, review another's protocol, reanalyze another's data, or recap another's literature summary and reach the same conclusions about the size of an effect. Clearly, as the role of judgment increases in any of these operations, the results become increasingly subjective. Typically, reproducibility should decrease (and subjectivity increase) to the extent that a problem

attracts scientists with diverse training or falls into a field that has yet to reach consensus on basic issues of methodology.

The second sense of objectivity means immune to the influence by value considerations. One's interpretations of data should not be biased by one's political views or pecuniary interests. Applied sciences naturally have developed great sensitivity to such problems and are able to invoke some penalties for detected violations. There is, however, little possibility of regulating the ways in which values influence other acts, such as one's choice of topics to study or ignore. Some of these choices might be socially sanctioned, in the sense that one's values are widely shared (e.g., deciding to study cancer because it is an important problem); other choices might be more personal (e.g., not studying an issue because one's employer does not wish to have troublesome data created on that topic). Although a commitment to separating issues of fact from issues of value is a fundamental aspect of intellectual hygiene, a complete separation is never possible (see Section III).

At times, this separation is not even desired—as when experts offer their views on how risks should be managed. Because they mix questions of fact and value, such views might be better thought of as the opinions of experts rather than as expert opinions, a term that should be reserved for expressions of substantive expertise. It would seem as though members of the public are the experts when it comes to striking the appropriate trade-offs between costs, risks, and benefits. That expertise is best tapped by surveys, hearings, and political campaigns.

Of course, there is no all-purpose public any more than there are all-purpose experts. The ideal expert on a matter of fact has studied that particular issue and is capable of rendering a properly qualified opinion in a form useful to decision makers. Using the same criteria for selecting value experts might lead one to philosophers, politicians, psychologists, sociologists, clergy, intervenors, pundits, shareholders, or well-selected bystanders. Thus, one might ask, "in what sense," whenever someone says "expert" or "public" (Schnaiburg, 1980; Thompson, 1980). This appendix uses "expert" in the restrictive sense and "public" or "laypeople" to refer to everyone else, including scientists in their private lives.

LAYPEOPLE AND EXPERTS ARE SPEAKING DIFFERENT LANGUAGES

Explicit risk analyses are a fairly new addition to the repertoire of intellectual enterprises. As a result, risk experts are only beginning to reach consensus on basic issues of terminology and methodology, such as how to define risk (see Section III). Their communications to the public reflect this instability. They are only beginning to express a sufficiently coherent perspective to help the public sort out the variety of meanings that "risk" could have. Under these circumstances some miscommunication may be inevitable. Studies (Slovic et al., 1979, 1980) have found that when expert risk assessors are asked to assess the risk of a technology on an undefined scale, they tend to respond with numbers that approximate the number of recorded or estimated fatalities in a typical year. When asked to estimate average year fatalities, laypeople produce fairly similar numbers. When asked to assess risk, however, laypeople produce quite different responses. These estimates seem to be an amalgam of their average-year fatality judgments, along with their appraisal of other features, such as a technology's catastrophic potential or how equitably its risks are distributed. These catastrophic potential judgments match those of the experts in some cases, but differ in others (e.g., nuclear power).

On semantic grounds, words can mean whatever a population group wants them to mean, as long as that usage is consistent and does not obscure important substantive differences. On policy grounds, the choice of a definition is a political question regarding what a society should be concerned about when dealing with risk. Whether we attach special importance to potential catastrophic losses of life or convert such losses to expected annual fatalities (i.e., multiply the potential loss by its annual probability of occurrence) and add them to the routine toll is a value question—as would be a decision to weight those routine losses equally rather than giving added weight to losses among the young (or among the nonbeneficiaries of a technology).

For other concepts that recur in risk discussions, the question of what they do or should mean is considerably murkier. It is often argued, for example, that different standards of stringency should apply to voluntarily and involuntarily incurred risks (e.g., Starr, 1969). Hence, for example, skiing could (or should) legitimately be a more hazardous enterprise than living below a major dam. Although there is general agreement among experts and laypeople about the voluntariness of food preservatives and skiing, other technologies are more problematic (Fischhoff et al., 1978b; Slovic et al., 1980). There is considerable disagreement within expert and lay groups in their ratings of the voluntariness of technologies such as prescription antibiotics, commercial aviation, handguns, and home appliances. These disagreements may reflect differences in the exposures considered; for example, use of commercial aviation may be voluntary for vacationers, but involuntary for certain business people (and scientists). Or, they may reflect disagreements about the nature of society or the meaning of the term. For example, each decision to ride in a car may be voluntarily undertaken and may, in principle, be foregone (i.e., by not traveling or by using an alternative mode of transportation); but in a modern industrial society, these alternatives may be somewhat fictitious. Indeed, in some social sets, skiing may be somewhat involuntary. Even if one makes a clearly volitional decision, some of the risks that one assumes may be indirectly and involuntarily imposed on one's family or the society that must pick up the pieces (e.g., pay for hospitalization due to skiing accidents).

Such definitional problems are not restricted to "social" terms such as "voluntary." Even a technical term such as "exposure" may be consensually defined for some hazards (e.g., medical x rays), but not for others (e.g., handguns). In such cases, the disagreements within expert and lay groups may be as large as those between them. For orderly debate to be possible, one needs some generally accepted definition for each important term—or at least a good translating dictionary. For debate to be useful, one needs an explicit analysis of whether each concept, so defined, makes a sensible basis for policy. Once they have been repeated often enough, ideas such as the importance of voluntariness or catastrophic potential tend to assume a life of their own. It does not go without saying that society should set a double standard on the basis of voluntariness or catastrophic potential, however they are defined.

LAYPEOPLE AND EXPERTS ARE SOLVING DIFFERENT PROBLEMS

Many debates turn on whether the risk associated with a particular configuration of a technology is acceptable. Although these disagreements may be interpreted as reflecting conflicting social values or confused individual values, closer examination suggests that the acceptable-risk question itself may be poorly formulated (Otway and von Winterfeldt, 1982).

To be precise, one does not accept risks—one accepts options that entail some level of risk among their consequences. Whenever the decision-making process has considered benefits or other (nonrisk) costs, the most acceptable option need not be the one with the least risk. Indeed, one might choose (or accept) the option with the highest risk if it had enough compensating benefits. The attractiveness of an option depends on its full set of relevant positive and negative consequences (Fischhoff, Lichtenstein, et al., 1981).

In this light, the term "acceptable risk" is ill defined unless the options and consequences to be considered are specified. Once the options and consequences are specified, "acceptable risk" might be used to denote the risk associated with the most acceptable alternative. When using that designation, it is important to remember its context dependence. That is, people may disagree about the acceptability of risks not only because they disagree about what those consequences are (i.e., they have different risk estimates) or because they disagree about how to evaluate the consequences (i.e., they have different values), but also because they disagree about what consequences and options should be considered.

Some familiar policy debates might be speculatively attributed, at least in part, to differing conceptions of what the set of possible options is. For example, saccharin (with its risks) may look unacceptable when compared with life without artificial sweeteners (one possible alternative option). Artificial sweeteners may, however, seem more palatable when the only alternative option considered is another sweetener that appears to be more costly and more risky. Or, nuclear power may seem acceptable when compared with alternative sources of generating electricity (with their risks and costs), but not so acceptable when aggressive conservation is added to the option set. Technical people from the nuclear industry seem to prefer the narrower problem definition, perhaps because they prefer to concentrate on the kinds of solutions most within their domain of expertise. Citizens involved in energy debates may feel themselves less narrowly bound; they may also be more comfortable with solutions, such as conservation, that require their kind of expertise (Bickerstaffe and Peace, 1980).

People who agree about the facts and share common values may still disagree about the acceptability of a technology because they have different notions about which of those values are relevant to a particular decision. For example, all parties may think that equity is a good thing in general, without agreeing also that energy policy is the proper arena for resolving inequities. For example, some may feel that both those new inequities caused by a technology and those old ones endemic to a society are best handled separately (e.g., through the courts or with income policies).

Thus, when laypeople and experts disagree about the acceptability of a risk, one must always consider the possibility that they are addressing different problems, with different sets of alternatives or different sets of relevant consequences. Assuming that each group has a full understanding of the implications of its favored problem definition, the choice among definitions is a political question. Unless a forum is provided for debating problem definitions, these concerns may emerge in more indirect ways (Stallen, 1980).

DEBATES OVER SUBSTANCE MAY DISGUISE BATTLES OVER FORM, AND VICE VERSA

In most political arenas, the conclusion of one battle often sets some of the initial conditions for its successor. Insofar as risk management decisions are shaping the economic and political future of a country, they are too important to be left to risk managers (Wynne, 1980). When people from outside the risk community enter risk battles, they may try to master the technical details or they may concentrate on monitoring and shaping the risk management process itself. The latter strategy may exploit their political expertise and keep them from being outclassed on technical issues. As a result, their concern about the magnitude of a risk may emerge in the form of carping about how it has been studied. They may be quick to criticize any risk assessment that does not have such features as eager peer review, ready acknowledgment of uncertainty, or easily accessible documentation. Even if they admit that these features are consonant with good research, scientists may resent being told by laypeople how to conduct their business even more than they resent being told by novices what various risks really are.

Lay activists' critiques of the risk assessment process may be no less irritating, but somewhat less readily ignored, when they focus on the way in which scientists' agendas are set. As veteran protagonists in hazard management struggles know, without scientific information it may be hard to arouse and sustain concern about an issue, to allay inappropriate fears, or to achieve enough certainty to justify action.

However, information is, by and large, created only if someone has a (professional, political, or economic) use for it. Whether the cause is fads or finances, failure to study particular topics can thwart particular parties and may lead them to impugn the scientific process.

At the other extreme, debates about political processes may underlie disputes that are ostensibly about scientific facts. As mentioned earlier, the definition of an acceptable-risk problem circumscribes the set of relevant facts, consequences, and options. This agenda setting is often so powerful that a decision has effectively been made once the definition is set. Indeed, the official definition of a problem may preclude advancing one's point of view in a balanced fashion. Consider, for example, an individual who is opposed to increased energy consumption but is asked only about which energy source to adopt. The answers to these narrower questions provide a de facto answer to the broader question of growth. Such an individual may have little choice but to fight dirty, engaging in unconstructive criticism, poking holes in analyses supporting other positions, or ridiculing opponents who adhere to the more narrow definition. This apparently irrational behavior can be attributed to the rational pursuit of officially unreasonable objectives.

Another source of deliberately unreasonable behavior arises when participants in technology debates are in it for the fight. Many approaches to determining acceptable-risk levels (e.g., cost–benefit analyses) make the political-ideological assumption that our society is sufficiently cohesive and common-goaled that its problems can be resolved by reason and without struggle. Although such a "get on with business" orientation will be pleasing to many, it will not satisfy all. For those who do not believe that society is in a fine-tuning stage, a technique that fails to mobilize public consciousness and involvement has little to recommend it. Their strategy may involve a calculated attack on what they interpret as narrowly defined rationality (Campen, 1985).

A variant on this theme occurs when participants will accept any process as long as it does not lead to a decision. Delay, per se, may be the goal of those who wish to preserve some status quo. These may be environmentalists who do not want a project to be begun or industrialists who do not want to be regulated. An effective way of thwarting practical decisions is to insist on the highest standards of scientific rigor.

LAYPEOPLE AND EXPERTS DISAGREE ABOUT WHAT IS FEASIBLE

Laypeople are often berated for misdirecting their efforts when they choose risk issues on which to focus their energies. However, a more careful diagnosis can often suggest several defensible strategies for setting priorities. For example, Zentner (1979) criticizes the public because its rate of concern about cancer (as measured by newspaper coverage) is increasing faster than the cancer rate. One reasonable explanation for this pattern is that people may believe that too little concern has been given to cancer in the past (e.g., our concern for acute hazards like traffic safety and infectious disease allowed cancer to creep up on us). A second is that people may realize that some forms of cancer are among the only major causes of death that experience increasing rates.

Systematic observation and questioning are, of course, needed to tell whether these speculations are accurate (and whether the assumption of rationality holds in this particular case). False positives in divining people's underlying rationality can be as deleterious as false negatives. Erroneously assuming that laypeople understand an issue may deny them a needed education; erroneously assuming that they do not understand may deny them a needed hearing. Pending systematic studies, these error rates are likely to be determined largely by the rationalist or emotionalist cast of one's view of human nature.

Without solid evidence to the contrary, perhaps the most reasonable general assumption is that people's investment in problems depends on their feelings of personal efficacy. That is, they are unlikely to get involved unless they feel that they can make a difference, personally or collectively. In this light, their decision-making process depends on a concern that is known to influence other psychological processes: perceived feelings of control (Seligman, 1975). As a result, people will deliberately ignore major problems if they see no possibility of effective action. Here are some reasons why they might reject a charge of "misplaced priorities" when they neglect a hazard that poses a large risk:

- the hazard is needed and has no substitutes;
- the hazard is needed and has only riskier substitutes;
- no feasible scientific study can yield a sufficiently clear and incontrovertible signal to legitimate action;

- the hazard is distributed naturally, and hence cannot be controlled;
- no one else is worried about the risk in question, and thus no one will heed messages of danger or be relieved by evidence of safety; and
- no one is empowered to or able to act on the basis of evidence about the risk.

Thus, the problems that actively concern people need not be those whose resolution they feel should rank highest on society's priorities. For example, one may acknowledge that the expected deaths from automobile accidents over the next century are far greater than those expected from nuclear power, and yet still be active only in fighting nuclear power out of the conviction, "Here, I can make a difference. This industry is on the ropes now. It's important to move in for the kill before it becomes as indispensable to American society as automobile transportation."

Thus, differing priorities between experts and laypeople may not reflect disagreements about the size of risks, but differing opinions on what can be done about them. At times, the technical knowledge or can-do perspective of the experts may lead them to see a broader range of feasible actions. At other times, laypeople may feel that they can exercise the political clout needed to make some options happen, whereas the experts feel constrained to doing what they are paid for. In still other cases, both groups may be silent about very large problems because they see no options.

LAYPEOPLE AND EXPERTS SEE THE FACTS DIFFERENTLY

There are, of course, situations in which disputes between laypeople and experts cannot be traced to disagreements about objectivity, terminology, problem definitions, process, or feasibility. Having eliminated those possibilities, one may assume the two groups really do see the facts of the matter differently. Here, it may be useful to distinguish between two types of situations: those in which laypeople have no source of information other than the experts, and those in which they do. The reasonableness of disagreements and the attendant policy implications look quite different in each case.

How might laypeople have no source of information other than the experts, and yet come to see the facts differently? One way is for the experts' messages not to get through intact, perhaps because: (1)

The experts are unconcerned about disseminating their knowledge or hesitant to do so because of its tentative nature; (2) only a biased portion of the experts' information gets out, particularly when the selection has been influenced by those interested in creating a particular impression; (3) the message gets garbled in transmission, perhaps due to ill-informed or sensationalist journalists; or (4) the message gets garbled upon reception, either because it was poorly explicated or because recipients lacked the technical knowledge needed to understand the message (Friedman, 1981; Hanley, 1980; Nelkin, 1977). For example, Lord Rothschild (1978) has noted that the BBC does not like to trouble its listeners with the confidence intervals surrounding technical estimates.

A second way of going astray is to misinterpret not the substance, but the process of the science. For example, unless an observer has reason to believe otherwise, it might seem sensible to assume that the amount of scientific attention paid to a risk is a good measure of its importance. Science can, however, be more complicated than that, with researchers going where the contracts, limelight, blue-ribbon panels, or juicy controversies are. In that light (and in hindsight), science may have done a disservice to public understanding by the excessive attention it paid to saccharin ("scientists wouldn't be so involved if this were not a major threat").

A second aspect of the scientific process that may cause confusion is its frequent disputatiousness. It may be all too easy for observers to feel that "if the experts can't agree, my guess may be as good as theirs" (Handler, 1980). Or, they may feel justified in picking the expert of their choice, perhaps on spurious grounds, such as assertiveness, eloquence, or political views. Indeed, it may seldom be the case that the distribution of lay opinions on an issue does not overlap some of the distribution of expert opinions. At the other extreme, laypeople may be baffled by the veil of qualifications that scientists often cast over their work. All too often, audiences may be swayed more by two-fisted debaters (eager to make definitive statements) than by two-handed scientists (saying "on the one hand X, on the other hand Y," in an effort to achieve balance).

In each of these cases, the misunderstanding is excusable, in the sense that it need not reflect poorly on the public's intelligence or on its ability to govern itself. It would, however, seem hard to justify using the public's view of the facts instead of or in addition to the experts' view. A more reasonable strategy would seem to be attempts at education. These attempts would be distinguished from attempts at propaganda by allowing for two-way communication, that is, by being open to the possibility that even when laypeople appear misinformed, they may still have defensible reasons for seeing things differently than do the experts.

For laypeople to disagree reasonably, they would have to have some independent source of knowledge. What might that be? One possibility is that they have a better overview on scientific debates than do the active participants. Laypeople may see the full range of expert opinions and hesitations, immune to the temptations or pressures that actual debaters might feel to fall into one camp and to discredit skeptics' opinions. In addition, laypeople may not feel bound by the generally accepted assumptions about the nature of the world and the validity of methodologies that every discipline adopts in order to go about its business. They may have been around long enough to note that many of the confident scientific beliefs of yesterday are confidently rejected today (Frankel, 1974). Such lay skepticism would suggest expanding the confidence intervals around the experts' best guess at the size of the risks.

Finally, there are situations in which the public, as a result of its life experiences, is privy to information that has escaped the experts (Brokensha et al., 1980). To take three examples: (1) The MacKenzie Valley Pipeline (or Berger) Inquiry discovered that natives of the far North knew things about the risks created by ice-pack movement and sea-bed scouring that were unknown to the pipeline's planners (Gamble, 1978); (2) postaccident analyses often reveal that the operators of machines were aware of problems that the designers of those machines had missed (Sheridan, 1980); and (3) scientists may shy away from studying behavioral or psychological effects (e.g., dizziness, tension) that are hard to measure, and yet still are quite apparent to the individuals who suffer from them. In such cases, lay perceptions of risk should influence the experts' risk estimates (Cotgrove, 1982; Wynne, 1983).

SUMMARY

It is tempting to view others in simplistic terms. Cognitively, one can save mental effort by relying on uncomplicated labels like "the hysterical public" or "the callous experts." Motivationally, properly chosen labels can affirm one's own legitimacy. By the same token, such interpretations can both obstruct the understanding of conflicts (by blurring significant distinctions) and hamper their resolution

(by bolstering self-serving characterizations). The following section begins by explaining the consequences of such stereotyping for risk communication by discussing the sort of communication strategies that can follow from simplistic interpretations of the controversy. It continues to outline principles for more complex strategies. These can inform both those designing communications programs and those receiving them.

V
STRATEGIES FOR RISK COMMUNICATION

CONCEPTS OF RISK COMMUNICATION

Risk communication is a collective noun for a variety of procedures expressing quite different attitudes toward the relationship between a society's laypeople and its technical-managerial elite (Covello et al., 1986). At one extreme lies the image of an inactive public docilely waiting for the transmission of vital information from those who know better. Within this perspective, the communication process involves a source, a channel, and a receiver (to use one set of technical terms common among social scientists). Although conceptually simple, this characterization still forces one to consider myriad details about each component. For example (Hovland et al., 1953): How well trusted is the source? Is it a corporate entity, capable of speaking with a single voice, or does it sometimes contradict itself? How much experience and language does the source share with the receivers? How much time does it have to prepare its messages? What are the legal restrictions on how much it can say?

At the other extreme lie highly interactive images of the communication process, in which the public shares responsibility for the social management of risks. Such processes, which require exchanges of information, could, in principle, be viewed as special cases of the source–channel–receiver model. However, using that model (and the research associated with it) requires bearing in mind the notion that these "receivers" are actively shaping the messages that they receive and perhaps even the research conducted in order to create the substance of those messages (Kasperson, 1986).

One way of diagnosing the nature of specific risk communication processes is in terms of the philosophies that guide those who design them. The following discussion describes some generic strategies in terms of their strengths and limitations. The discussion after that considers some more integrative design principles. Together, they are intended to create a framework for responsibly using the more technical material on communication design presented in the final section. That material assumes an understanding of the role of information in the risk management (including communication) process (Johnson and Covello, 1987; Rayner and Cantor, 1987).

SOME SIMPLE STRATEGIES

The technical and policy issues involved in making risk management decisions are complex enough in themselves. Dealing with public perceptions of risks creates an additional level of complexity for risk managers. One possible response to this complexity is to look for some "quick fix" that will deal with the public's needs. Unfortunately for the risk manager, these strategies are both hard to execute well by themselves and unlikely to be sufficient even if they are well executed. At times, these simple solutions seem to reflect a deep misunderstanding of the public's role in risk management, reflecting perhaps a belief that the human element in risk management can be engineered in the same way as mechanical and electronic elements. Undertaken in isolation and with these unrealistic expectations, such strategies can produce mutually frustrating communication programs. The following are some of the more common of these simple strategies for dealing with risk controversies, presented in caricature form to highlight their underlying motivations and inherent limitations.

Give the Public the Facts

The assumption underlying this strategy is that if laypeople only knew as much as the experts, they would respond to hazards in the same way. Undertaken insensitively, this strategy can result in an incomprehensible deluge of technical details, telling the public more than it needs to know about specific risk research results, and much less than it needs to know about the quality of the research (and about how to make the decisions that weigh most heavily on its mind). Concentrating communications on the transmission of information also ignores the possibility that there are legitimate differences between the public and the experts regarding either the goals or the facts of risk management.

Sell the Public the Facts

The premise here is that the public needs persuasion, rather than education. It often follows the failure of an information campaign to win public acceptance for a technology. Undertaken heavy-handedly, this approach may amount to little more than repeating more loudly (or fancily) messages that the public has already rejected. Here, as elsewhere, obvious attempts at manipulation can breed resentment.

Give the Public More of What It Has Gotten in the Past

The underlying assumption here is that the public will accept in the future the kinds of risks that it has accepted in the past. If true, then what the public wants (and will accept) can be determined simply by examining statistics showing the risk–benefit trade-offs involved in existing technologies. This "revealed preference" philosophy ignores the fact, consistently revealed by opinion polls showing great public support for environmental regulations, that people are unhappy with how risks have been managed in the past. The risks that people have tolerated are not necessarily acceptable to them. As a result, giving them more of the same means enshrining past inequities in future decisions. In principle, this approach attaches no importance to educating the public, to creating a constituency for risk policies, or to involving the public in the political process. It seems to respect the public's wishes, while keeping the public itself at arm's length.

Give the Public Clear-Cut, Noncontroversial Statements of Regulatory Philosophy

The assumption underlying this family of approaches is that people do not want facts, but instead the assurance that they are being protected. That is, whatever the risks may be, they are in line with government policy. Examples in the United States include the Delaney clause, prohibiting carcinogenic additives in foods, and the Nuclear Regulatory Commission's "safety goals for nuclear power," describing how risky it will allow the technology to be. Each policy is stated in terms of levels of acceptable risk, as though laypeople are too unsophisticated to understand, in the context of technology management, the sort of risk–benefit trade-offs that they routinely make in everyday life, such as when they undergo medical treatments or pursue hazardous occupations. Moreover, such simple statements provide little guidance for many real situations—by denying the complexity of the (risk–benefit) decisions that needed to be made. If perceived as hollow, then they will do little to reassure the public.

Let the Marketplace Decide

Another hope for risk communication is that risks will be understood when communicated in the context of specific consumer

decisions. One variant on this approach is the claim that reducing government regulation will allow people to decide independently what risks they are willing to accept, with the courts addressing any excesses. A second variant is providing quantitative risk information along with goods and drugs. It makes optimistic assumptions regarding laypeople's ability to know enough to fend for themselves with all life's risks. The assumption of personal responsibility and the motivation to get it right are meant to prompt efficient acquisition and understanding. It assumes that people will recognize the limits to their risk perceptions and grasp the risk information presented to them. A threat to any approach emphasizing self-reliance is that people might not want to defend their own welfare when it comes to health and safety, especially where risks have long latencies and it is impossible to prove the source of a health risk (and obtain redress).

Put Risk Managers on the Firing Line

The assumption underlying this strategy is that what the public needs in order to understand risk issues is a coherent story from a single credible source. Examples might include the Nuclear Regulatory Commission's reliance on a single spokesperson as the Three Mile Island incident wore on and the assumption of center stage by the president of Union Carbide after the chemical gas leak in Bhopal, India. This strategy can reduce the confusion created by incomplete conflicting messages, although only if the manager has good communication skills or is sensitive to listeners' information needs; that is, there must be both substance and style. Oversimplifications, misrepresentations, and unacceptable policies are just that, even if they come from a nice guy. This approach can also create a bottleneck for understanding the public's concerns to the extent that the single source of information must also be the single recipient.

Involve Local Communities in Resolving Their Own Risk Management Problems

This approach assumes that people will be flexible and realistic about trade-offs when they see—and have responsibility for—the big picture. Such an approach can founder when the community lacks real decision-making authority or the technical ability to understand its alternatives. It may also founder when those alternatives accept perceived past inequities (e.g., reduce chronic poverty by accepting a hazardous waste dump) or are of the jobs-versus-health variety that people expect government to help them resolve. Ensuring the informed consent of the governed for the risks to which they are exposed is a laudable goal. However, its achievement requires that people have tolerable choices, adequate information, and the ability to identify which course of action is in their own best interests.

CONCEPTUALIZING COMMUNICATION PROGRAMS

Despite their flaws, these simple strategies all have some merit. It is important to give people the facts and to be persuasive when the facts do not speak for themselves or when existing prejudices must be overcome. It is also important to maintain some consistency with past risk management decisions, to expound clear policies, to exploit the wisdom of the marketplace, to encourage direct communication between risk managers and the public, and to give communities meaningful control over their own destinies. The problem is that each strategy oversimplifies the nature of risk issues and the public's involvement with them. When risk managers pin unrealistic hopes on such strategies, then the opportunity to address the public's needs more comprehensively is lost. When these hopes are not met, the frustration that follows is often directed at the public.

It is both unfair and corrosive for the social fabric to criticize laypeople for responding inappropriately to risk situations for which they were not adequately prepared. It is tragic and dangerous when members of our technical elite feel that they have devoted their lives to creating a useful technology (e.g., nuclear power) only to have it rejected by a foolish and unsophisticated public. Likewise, it is painful and unfortunate when the public labels those elites as evil and arrogant.

Risk management requires allocating resources and making trade-offs between costs and benefits. Thus, it inherently involves conflicts. Both the substance and the legitimacy of these conflicts are obscured, however, when the participants come to view them as struggles between the forces of good and evil, or of wisdom and stupidity. Effective solutions will have to be respectful solutions, recognizing both the legitimacy and complexity of the public's perspective, giving it no more and no less credit for reasonableness than it deserves.

How can the preceding observations about risk perceptions (and the research literature from which they were drawn) be used to design better procedures for dealing with risk controversies?

One necessary starting point is a detailed consideration of the nature of the risk that the public must understand. That consideration must cover not only the best available technical estimates for the magnitude of the risk, but also the best available psychological evidence on how people respond to that kind of risk. Research has shown, for example, that people have special demands for safety—and reassurance—when risks are perceived to have delayed effects or catastrophic potential, and when risks appear to be poorly understood or out of people's personal control (Slovic, 1986; Vlek and Stallen, 1980, 1981; von Winterfeldt et al., 1981). Such risks are likely to grab people's attention and create unrest until they can be put in some acceptable perspective. They demand greater communication resources, with particular attention devoted to creating an atmosphere of trust. Perhaps paradoxically, people may need to be treated with the greatest respect in those situations in which they may seem most emotional (or most human) (Eiser, 1982; Weinstein, 1987).

A second necessary starting point is a detailed description of how information about risk can reach people (Johnson and Covello, 1987; Rubin and Sachs, 1973; Schudson, 1978). Such information may be the result of accidents at various distances away and attributed to various causes (e.g., malfunctions, human error, sabotage) or of mere "incidents," such as newspaper exposés, siting controversies, false alarms, or government inquiries. Proactively, this analysis will show the opportunities for reaching people. For example, is there a chance to educate at least some of the public in advance, or can one only prepare materials for times of crisis? Reactively, this analysis should help one anticipate what people will already know (or believe) when the time comes for systematic communication. It may show that people are buffeted by confusing, contradictory, and erroneous messages—or that they have some basic understanding within which they can integrate new information. In any case, communication must build on people's current mental representation of the technology—even if its first step is to challenge inappropriate beliefs and enhance people's ability to examine future information more critically.

Knowing what people do know allows a systematic analysis of what they need to know—the next point of departure in communicating with the public. In some cases, crude estimates of a technology's risks and benefits may be enough; in other cases, it may be important

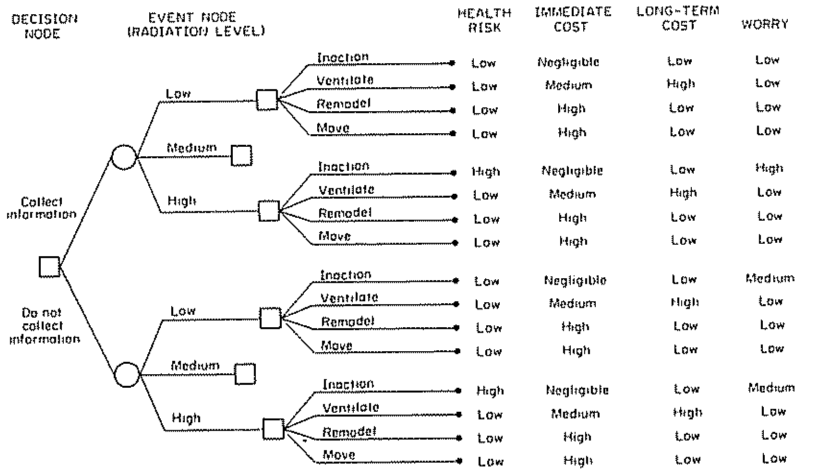

FIGURE V.1 The radiation hazard in homes from the residents' perspective. SOURCE: Svenson and Fischhoff, 1985.

to know how a technology operates. The needs depend on the problems that the public is trying to solve: what to do in an emergency; how to react in a siting controversy; whether to eat vegetables, or whether to let their children do so; and so on. Perhaps the most efficient description would be in the terms of decision theory, such as the simple decision tree in Figure V.1, depicting the situation faced by the head of a household deciding whether to test for domestic radon accumulations. Such descriptions allow one to determine how sensitive these decisions are to different kinds of information, so that communication can focus on the things that people really need to know.

Producing comparable descriptions for the different actors in a risk management episode will help clarify sources of disagreement among them. Often the risk managers' decision problem (e.g., whether to ban EDB) will be quite different from the public's decision problem (e.g., whether to use blueberry muffin mix). For example, Figure V.2 shows the key decision problem that might face risk managers concerned about radon: what standard to set as expressing a tolerable level of exposure. The critical outcomes of this decision are quite different from those associated with the residents' focal decision of whether to test their homes for radon (Figure V.1). Failure

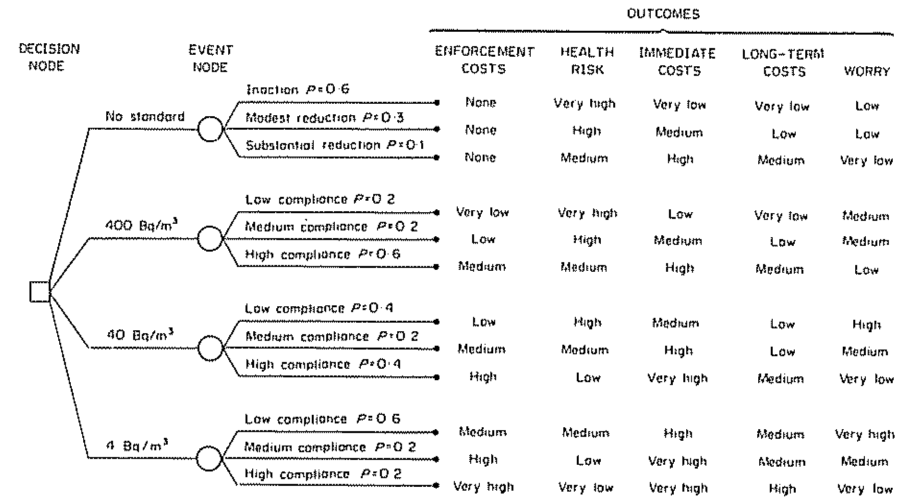

FIGURE V.2 The radiation hazard in homes from the authorities' perspective. SOURCE: Svenson and Fischhoff, 1985.

to address the public's information needs is likely to leave them frustrated and hostile. Failure to address the managers' own problems is likely to leave their eventual actions inscrutable. For telling their own story, the managers need a protocol that will ensure that all of the relevant parts get out, including what options they are legally allowed to consider, how they see the facts, and what they consider to be the public interest. Such comprehensive accounts are often absent from the managers' public pronouncements, preventing the public from responding responsibly and suggesting that the managers failed to consider the issues fully. The procedures offered in Section II as ways for the public (or the media) to discover what risk issues are all about might also be used proactively as ways to tell the public (or the media) directly about those risks.

After determining what needs to be said, risk managers can start worrying about how to say it. A common worry is that the public will not be able to understand the technical details of how a technology operates. Where those details are really pertinent, the services of good science writers and educators may be needed. Perhaps a more common problem is making the basic concepts of risk management clear. Just what is a one-in-a-million chance? What does it mean to protect wastes for a hundred generations? Must we inevitably set a value on human life when resources are allocated for risk reduction?

The psychological research described above has shown the difficulty of these concepts; it is beginning to show ways to communicate them meaningfully. The research base for addressing these obstacles to understanding is described in the next section.

Adopting such a deliberative approach to characterizing people's needs would help avoid the inadvertent insensitivity found in the Institute of Medicine's (1986) report, *Confronting AIDS*. The report noted, somewhat despairingly, that only 41 percent of the general public knew that AIDS was caused by a virus. Yet, although this fact is elemental knowledge for medical researchers, it has relatively little practical importance for laypeople—in the sense that one would be hard pressed to think of any real decision whose resolution hinged on knowing that AIDS was a virus. Laypeople interested in a deep understanding of the AIDS problem ought to know this fact. However, it is irrelevant to laypeople satisfied just to make reasonable decisions regarding AIDS. Such insensitivity is socially damaging insofar as it demeans the public in the eyes of the experts and prompts the provision of seemingly irrelevant communications.

Another example of this insensitivity to the needs of message recipients can be found in the advice literature about sexual assault (Morgan, 1986). Much of the research is performed and communicated without consideration for women's decision-making needs (Furby and Fischhoff, in press). Most studies concentrate on significance levels, whereas what women need is reliable information on effect size. That is, women need to know not only whether a strategy makes a difference, but how much of a difference. A second form of insensitivity to women's decision-making needs is that few studies collect data on the temporal order of strategies and consequences. As a result, although if greater physical resistance by women were associated with greater violence by men, one would not know which causes which. A third form of insensitivity can be found in recommendations telling women how to respond to different kinds of assailants, without considering whether women can even make such diagnoses under real-life conditions or without reporting the overall prevalence (or "base rates") of the different assailant types, an essential piece of information for making any diagnosis. Finally, some studies actually made the "base-rate fallacy" (Bar-Hillel, 1980; Kahneman and Tversky, 1972), concluding, say, that screaming is more effective than fighting because, among women who escape, 80 percent do the former and only 20 percent do the latter.

Taking the details of risk perceptions seriously means reconciling ourselves to a messy process. In managing risks, society as a whole is slowly and painfully learning how to make deliberative decisions about very difficult issues. Avoiding frustration with the failures and with the public that seems responsible for them will help us keep the mental health and mutual respect needed to get through it all.

EVALUATING COMMUNICATION PROGRAMS

Testing Risky Treatments

If they were creating risks rather than explaining them, risk communicators would be subject to various political, legal, and social constraints. If the treatment involved a medical intervention, then there would be a comparable tangle of restrictions. What analogous responsibilities are incumbent on those who treat others with information?

A minimal requirement might be that a communication have positive expected value. That is, its anticipated net effect should be for the good, considering the magnitude and likelihood of possible consequences. Releasing a communication program that flunked this test would be like authorizing a drug with uncompensated side effects.

A minimal standard of proof for passing this minimal test is expert judgment. Thus, a communication technique could be approved if it were "generally regarded as safe" and seemed likely to be at least somewhat effective. Such reliance on experts' intuitions creates the same discomfort as comparable proposals for grandparenting existing drugs or additives because they are familiar and appear to be safe. How do we know they work? Might negative effects simply have escaped notice or measure? Just what do these experts know? Can they be trusted?

More convincing would be empirical evidence from a basic science of risk communication providing some a priori basis for predicting the effects of particular communications. That evidence could be positive, showing that a communication draws on a demonstrated cognitive ability [e.g., people can understand quantitative probabilities, as long as they are not too small (Beyth-Marom, 1982)]. Or, it could be negative, showing that a communication demands a kind of understanding that is not widely distributed [e.g., people have trouble realizing how the probability of failure accumulates from repeated events, such as using a contraceptive device or being exposed to a disease (Bar-Hillel, 1973)].

More convincing still is evidence from a test of the communication itself, performed with individuals like its ultimate recipients and in a setting like that in which it will ultimately be administered. If that setting must be simulated, then the simulation should capture both those features of the actual communication context that interfere with understanding (e.g., talking to friends during the transmission) and those features that can enhance comprehension (e.g., discussing the transmission with friends) (Turner and Martin, 1985).

Evaluative Criteria

Performing an evaluation requires a clear, operable definition of the consequences to be desired and avoided. With medical treatments, identifying the consequences is usually a straightforward process—they are various possible health effects, some good and some bad. What might be more complicated is measuring some of the effects (e.g., those involving delayed consequences) and determining their relative importance. Although medical personnel and their clients are likely to agree about which outcomes are good and which are bad, they need not agree about how good and how bad the outcomes are. For example, they might feel differently about trade-offs between short- and long-term effects or between changes in quality of life and in expected longevity (McNeil et al., 1978). As a result, even after a definitive evaluation, there may be no universal recommendation. A well-understood treatment might be right for some people, but wrong for others.

In evaluating communication programs, similar issues arise, although with a few additional wrinkles. Potential consequences must still be identified. However, the set seems less clearly defined. There are the good and bad health effects, but they may be hard to observe. If a communication causes undue concern, then there may be stress-related effects, but they tend to be quite diffuse (e.g., a few more cases of child abuse, depression, divorce, and so on, scattered through the treated population) (Elliot and Eisdorfer, 1982). On the other side of the ledger, if people do engage in health-enhancing behavior, then the influence of the focal communication must be isolated from that of other information sources (including, perhaps, continued rumination about an issue).

Difficulties in observing the effects of ultimate interest may divert attention to more observable effects closer to the treatment.

One possibility that arises with communication programs (unlike conventional medical treatments) is assessing comprehension of the message. If people have not understood the message, then an appropriate response seems unlikely. The simplest test of comprehension might be remembering the facts of a message. Those recipients who pass it would, however, still have to be tested for whether they are able to use those remembered facts in their decision making. Those who fail the test would still have to be tested for whether they have heard the message, but chose to reject it. Rejection might mean distrusting the source's competence or its motives. That is, the communicators may not seem to know what they are talking about or they may seem inadequately concerned about the recipients' welfare.

Setting Objectives for Communication Programs

It is accepted wisdom that program planning of any sort ought to begin with an explicit statement of objectives, in the light of which a program's elements can be selected and its effects evaluated. Figure V.3 offers one conceptualization of risk communication programs, categorized according to their primary objective.

According to Covello et al. (1986:172–173):

> In the real world, these four types of risk communication tasks overlap substantially, but they still can be conceptually differentiated. The task of informing and educating the public can be considered primarily a non-directive, although purposeful, activity aimed at providing the lay public with useful and enlightening information. In contrast, both the task of encouraging behavior change and personal protective action and that of providing disaster warnings and emergency information can be considered primarily directive activities aimed at motivating people to take specific types of action. These three tasks, in turn, differ from the task of involving individuals and groups in joint problem solving and conflict resolution, in which officials and citizens exchange information and work together to solve health and environmental problems.

As can be seen from Figure V.3, much risk communication is initiated with the communicators' benefit foremost in mind. For example, the sponsors of a technology may wish to reassure a recalcitrant and alarmed public about its safety. If the public's worry is really unwarranted, then everyone comes out ahead: The technology will get a fairer shake and the public will be relieved of an unnecessary worry. The crucial question is what constitutes "unwarranted" concern. One possible definition is exaggerating the magnitude of the risk (or underestimating the magnitude of accompanying benefits).

TYPE 1: **Information and Education**

- Informing and educating people about risks and risk assessment in general.

  EXAMPLE: statistical comparisons of the risks of different energy production technologies.

TYPE 2: **Behavior Change and Protective Action**

- Encouraging personal risk-reduction behavior.

  EXAMPLE: advertisements encouraging people to wear seat belts.

TYPE 3: **Disaster Warnings and Emergency Information**

- Providing direction and behavioral guidance in disasters and emergencies.

  EXAMPLE: sirens indicating the accidental release of toxic gas from a chemical plant.

TYPE 4: **Joint Problem Solving and Conflict Resolution**

- Involving the public in risk management decision-making and in resolving health, safety, and environmental controversies

  EXAMPLE: public meetings about a possible hazardous waste site

FIGURE V.3 A typology of risk communication objectives. SOURCE: Covello et al., 1986.

In such cases, straight information messages might help. However, they need to be designed with an eye to implicit as well as explicit content. For example, if they are perceived as insistently repeating that "the risk is only X" (or that "the benefit is really Y"), then recipients may read between the lines, "and that ought to be good enough for you." Communicators may convince themselves about the rectitude of such implicit messages, feeling that expert knowledge about the size of risks generalizes to expert knowledge about their acceptability.

Certainly, people should be better off with better information. However, even well-informed people may dislike a technology if they feel that its benefits (to them) are not commensurate with its risks (to them), or that those benefits are substantially lower than the benefits enjoyed by a technology's sponsor. Honest communications should help people reach such determinations. As a result, neither the senders nor the recipients of messages should be faulted if more information leads to more opposition.

An alternative definition of "unwarranted concern" is "larger than the concern associated with hazards having equivalent risk." In more sophisticated versions, the comparison might be with concern over hazards having an equivalent relationship between risks and benefits. A popular contribution to the risk literature a decade ago was lists of disparate risks, chosen so that most were, arguably, accepted by most people (Cohen and Lee, 1979; Crouch and Wilson, 1982). The lists would also contain some favored technology (e.g., nuclear power) that should seemingly be accepted, by whatever criterion led to the acceptance of the other risks in the list. Such lists might, if thoughtfully assembled, help to educate readers' intuitions about the relative magnitude of different risks and the nature of very small risks (e.g., 10^{-6}), such as often appear in such lists. However, even recipients who accept the general idea of consistency that underlies such claims need not accept the particular form of consistency implied by the list (Covello et al., 1988). They may not endorse the particular definition of risk used in the list; they may not feel that all currently accepted (or tolerated or endured) risks are actually acceptable (in the sense that they have agreed voluntarily to the hazards bearing those risks and would not want lesser risks if those were available at a reasonable price). Nor need people accept even the weaker consistency claim that they should not worry more about any hazard than they worry about hazards that they believe to have greater risks. Section III discusses some of people's reasons for ignoring admittedly large hazards.

Comprehension of risk messages is seldom the consequence that is ultimately of interest. Rather, it is a potentially observable surrogate for actual improvements in well-being. A step closer to that consequence would be evidence that recipients of a message had connected their perception of its contents with the course(s) of action in their own best interests (i.e., what a decision theorist would prescribe, given recipients' definition of the situation). For achieving this goal, recipients could be left to their own devices, or they might be provided some help in connecting their beliefs and values with possible actions.

Assuming that it can be done in a neutral (noncoercive) way, providing such help changes the nature of the relationship. Rather than one party administering an informational treatment to another, the treater becomes more of an aide and servant. One particular expression of the change emerges in situations in which a communicator wishes to claim that people have given "informed consent" to the risks described in a communication (P.S. Appelbaum et al., 1987). That claim should interest people exposed to the risks only if it changes their bargaining position vis-à-vis the creator of the risks (e.g., "what's it worth to you for me to sign this release?" or "does that mean that I can force you to give me more information about potential adverse health effects?"). What people should care about is identifying the best choice of action. A communication serves that end if it provides people with the information that they need in a form that they can use. In this light, informed consent may be claimed when people have chosen the best possible course of action for themselves.

These criteria for evaluating risk communication, like those typically invoked for evaluating medical treatments, are focused on direct effects of simple interventions. However, any treatment is but one in a series (at least for those who survive). For example, treatment with an antibiotic might cause no immediate adverse side effects, but might still create an allergic condition that reduces the set of possible treatments for future maladies. Good communication can enhance recipients' actual and perceived ability to understand a risky world and deal with it effectively. Poor communication can do the opposite, reducing recipients' confidence in their own competence to manage the risks in their lives. Just as emotional involvement can impair understanding of the content of messages, so can misunderstanding messages produce unproductive emotions.

Institutional Controls

If risk communications were viewed as treatments, then they might also "enjoy" an institutional context like that created for medical treatments. One component might be review panels to scrutinize the protocols for testing or running communication programs. Such panels might both ensure that programs use suitable evaluation criteria (e.g., reflecting both senders' and recipients' needs) and examine messages for attempts to coerce or misinform. Review panels might also provide guidance on ethical issues. For example, if there is a commonly accepted "best" way to convey a certain kind of information, can one legitimately substitute new, experimental methods? How would that decision change as a function of the kind of testing that the accepted method had undergone? Or, what should be done with messages telling people that they are powerless to affect their fate (e.g., they have been exposed to a carcinogen with irreversible

effects, such as asbestos)? Recipients' natural concern over the risk could be aggravated by the feeling of helplessness, especially if the risk is perceived as having been imposed by someone else without providing proper consent or compensation. Do senders have a responsibility to provide counseling for those upset by their messages? Might they even restrict dissemination? How would the decision about the communication process change if the information would help recipients (or others) to mobilize their resources in responding to other hazards? If there are only limited resources for communication, who should receive them (e.g., those at greatest risk, those most responsive to available communication techniques, or those most accessible)?

The institutional context for medical treatments attempts not only to ensure that they are delivered properly, but also to address possible failures. Lists of counterindications accompany many treatments. Physicians are always on stand-by, ready to ameliorate the side effects of their treatments. Various mechanisms exist for collecting and disseminating (good and bad) experiences, for both veteran and experimental treatments. When the rate of side effects is unacceptable, either for a treatment or for a treater, government and professional bodies may stop the exposure. In the background of all these efforts to manage risks lurks the threat of legal proceedings to rectify unmanaged problems (e.g., malpractice and product liability suits). People are more likely to behave well when there are strong social norms for doing so and significant penalties for failure. The desire to be fair to all parties prompts a sharpening of standards.

It took many years to evolve these institutions and standards (many centuries, if one reaches back to Hippocrates). Judging by the various contemporary crises (e.g., malpractice, cost containment), they are still far from perfect. However, those imperfections pale before those of treatments with no such infrastructure. In cases in which an institutional context is created anew for a particular cause, it may be hard to get this degree of balance. For example, right-to-know laws have recently been enacted to ensure that workers receive information about occupational hazards. The laws are intended to help workers protect themselves on the job and to help employers protect themselves in court (by strengthening their claim that workers have given informed consent to bearing the risks). The criteria for evaluating these efforts seem to concentrate more on what is said than on what is understood, raising the threat of overloaded and overly technical messages filling the letter but not the intent of the law. The existence of such threats suggests a tenuous state of affairs for even the more developed areas of risk communication.

SUMMARY

Risk information is an important part of many human activities. Yet it is at most but a part. Understanding its role is essential to giving risk communication programs their basic shape, with appropriate objectives and realistic expectations. Such an analysis can help communicators avoid simplistic strategies that leave recipients, at best, unsatisfied and, at worst, offended by the failure to address their perceived needs. In some cases, these will be for better information; in other cases, they will be for better protection. Only after communication programs are recipient centered in this respect can they productively begin to be recipient centered in the sense of the following section, considering laypeople's strengths and weaknesses in understanding risk information.

VI PSYCHOLOGICAL PRINCIPLES IN COMMUNICATION DESIGN

Whenever they read a brochure, talk to their neighbors, or observe ominous activities at a local plant in order to understand the risks of a technology, people must rely on the same basic cognitive processes that they use to understand other events in their lives. As mentioned in Section II, the study of such processes is an involved pursuit, with many methodological nuances (like most sciences). To provide some access to the substantive results of such research, here are a number of relatively simple and generally supported statements about behavior. The difficulty in applying them to the prediction of real-life behavior is that life's situations are complex, meaning that various simple behaviors interact in ways that require a subtle analysis to understand.

PEOPLE SIMPLIFY

Most substantive decisions require people to deal with more nuances and details then they can readily handle at any one time. People have to juggle a multitude of facts and values when deciding, for example, whether to change jobs, trust merchants, or protest a toxic landfill. To cope with this information overload, people simplify. Rather than attempting to think their way through to comprehensive, analytical solutions to decision-making problems, people try to rely on habit, tradition, the advice of neighbors (or the media), and on general rules of thumb (e.g., nothing ventured, nothing gained). Rather than consider the extent to which human behavior varies from situation to situation, people describe other people in terms of all-encompassing personality traits, such as being honest, happy, or risk seeking (Nisbett and Ross, 1980). Rather than think precisely about the probabilities of future events, people rely on vague quantifiers, such as "likely" or "not worth worrying about"—terms that are also used differently by different people and by the same individual in different contexts (Beyth-Marom, 1982).

The same desire for simplicity can be observed when people press risk managers to categorize technologies, foods, or drugs as "safe" or "unsafe," rather than treating safety as a continuous variable. It can be seen when people demand convincing proof from scientists who can provide only tentative findings. It can be seen when people attempt to divide the participants in risk disputes into good guys and bad guys, rather than viewing them as people who, like themselves, have complex and interacting motives. Although such simplifications help people cope with life's complexities, they can also obscure the fact that most risk decisions involve gambling with people's health, safety, and economic well-being in arenas with diverse actors and shifting alliances.

ONCE PEOPLE'S MINDS ARE MADE UP, IT IS DIFFICULT TO CHANGE THEM

People are extraordinarily adept at maintaining faith in their current beliefs unless confronted with concentrated and overwhelming evidence to the contrary. Although it is tempting to attribute this steadfastness to pure stubbornness, psychological research suggests that some more complex and benign processes are at work (Nisbett and Ross, 1980).

One psychological process that helps people maintain their current beliefs is feeling little need to look actively for contrary evidence. Why look, if one does not expect that evidence to be very substantial or persuasive? For example, how many environmentalists read *Forbes* and how many industrialists read the Sierra Club's *Bulletin* in order to learn something about risks (as opposed to reading these publications to anticipate the tactics of an opposing side)? A second contributing thought process is the tendency to exploit the uncertainty surrounding apparently contradictory information in order to interpret it as being consistent with existing beliefs. In risk debates, a stylized expression of this proficiency is finding just enough problems with contrary evidence to reject it as inconclusive.

A third thought process that contributes to maintaining current beliefs can be found in people's reluctance to recognize when information is ambiguous. For example, the incident at Three Mile Island would have strengthened the resolve of any antinuclear activist who asked only, "how likely is such an accident, given a fundamentally unsafe technology?", just as it would have strengthened the resolve of any pronuclear activist who asked only, "how likely is the containment of such an incident, given a fundamentally safe technology?" Although a very significant event, Three Mile Island may not have revealed very much about the riskiness of nuclear technology as a whole. Nonetheless, it helped the opposing sides polarize their views. Similar polarization has followed the accident at Chernobyl, with opponents pointing to the "consequences of a nuclear accident" (which come with any commitment to nuclear power) and proponents pointing to the unique features of that particular accident (which are unlikely to be repeated elsewhere, especially considering the precautions instituted in its wake) (Krohn and Weingart, 1987).

PEOPLE REMEMBER WHAT THEY SEE

Fortunately, given their need to simplify, people are quite good at observing those events that come to their attention (and that they are motivated to understand) (Hasher and Zacks, 1984; Peterson and Beach, 1967). As a result, if the appropriate facts reach people in a responsible and comprehensible form before their minds are made up, there is a decent chance that their first impression will be the correct one. For example, most people's primary sources of information about risks are what they see in the news media and observe in their everyday lives. Consequently, people's estimates of the principal causes of death are strongly related to the number of people they know who have suffered those misfortunes and the amount of media coverage devoted to them (Lichtenstein et al., 1978).

Unfortunately for their risk perceptions (although fortunately for their well-being), most people have little firsthand knowledge of hazardous technologies. Rather, what laypeople see most directly are the outward manifestations of the risk management process, such as hearings before regulatory bodies or statements made by scientists to the news media. In many cases, these outward signs are not very reassuring. Often, they reveal acrimonious disputes between supposedly reputable experts, accusations that scientific findings have been distorted to suit their sponsors, and confident assertions that are disproven by subsequent research (Dietz and Rycroft, 1987; MacLean, 1987; Rothman and Lichter, 1987).

PEOPLE CANNOT READILY DETECT OMISSIONS IN THE EVIDENCE THEY RECEIVE

Not all problems with information about risk are as readily observable as blatant lies or unreasonable scientific hubris. Often, the information that reaches the public is true, but only part of the truth. Detecting such systematic omissions proves to be quite difficult (Tversky and Kahneman, 1973). For example, most young people know relatively few people suffering from the diseases of old age; nor are they likely to see those maladies cited as the cause of death in newspaper obituaries. As a result, young people tend to underestimate the frequency of these causes of death, while overestimating the frequency of vividly reported causes, such as murder, accidents, and tornadoes (Lichtenstein et al., 1978).

Laypeople are even more vulnerable when they have no way of knowing about information because it has not been disseminated. In principle, for example, patients could always ask their physicians whether they have neglected to mention any side effects of the drugs they prescribe. Likewise, people could always ask merchants whether there are any special precautions for using a new power tool, or ask proponents of a hazardous facility if their risk assessments have considered operator error and sabotage. In practice, however, these questions about omissions are rarely asked. It takes an unusual turn of mind to recognize one's own ignorance and insist that it be addressed.

As a result of this insensitivity to omissions, people's risk perceptions can be manipulated in the short run by selective presentation. Not only will people not know what they have not been told, but they will not even notice how much has been left out (Fischhoff et al., 1978a). What happens in the long run depends on whether the unmentioned risks are revealed by experience or by other sources of information. When deliberate omissions are detected, the responsible party is likely to lose all credibility. Once a shadow of doubt has been cast, it is hard to erase.

PEOPLE MAY DISAGREE MORE ABOUT WHAT RISK IS THAN ABOUT HOW LARGE IT IS

Given this mixture of strengths and weaknesses in the psychological processes that generate people's risk perceptions, there is no simple answer to the question "how much do people know and understand?" The answer depends on the risks and on the opportunities that people have to learn about them.

One obstacle to determining what people know about specific risks is disagreement about the definition of risk. (See Sections II and III for more complete discussions of different possible definitions of risk and other terms.) If laypeople and risk managers use the term risk differently, then they can agree on the facts about a specific technology but still disagree about its degree of riskiness. Several years ago, the idea circulated in the nuclear power industry that the

public cared much more about multiple deaths from large accidents than about equivalent numbers of casualties resulting from a series of small accidents. If this assumption were valid, then the industry would be strongly motivated to remove the threat of such large accidents. If removing the threat proved impossible, then the industry could argue that a death is a death and that in formulating social policy it is totals that matter, not whether deaths occur singly or collectively.

There were never any empirical studies to determine whether this was really how the public defined risk. Subsequent studies, though, have suggested that what bothers people about catastrophic accidents is the perception that a technology capable of producing such accidents cannot be very well understood or controlled (Slovic et al., 1984). From an ethical point of view, worrying about the uncertainties surrounding a new and complex technology such as nuclear power is quite a different matter than caring about whether a fixed number of lives are lost in one large accident rather than in many small accidents.

PEOPLE HAVE DIFFICULTY DETECTING INCONSISTENCIES IN RISK DISPUTES

Despite their frequent intensity, risk debates are typically conducted at a distance (Hance et al., 1988; Mazur, 1973). The disputing parties operate within self-contained communities and talk principally to themselves. Opponents are seen primarily through their writing or their posturing at public events. Thus, there is little opportunity for the sort of subtle probing needed to discover basic differences in how the protagonists think about important issues, such as the meaning of key terms or the credibility of expert testimony. As a result, it is easy to misdiagnose one another's beliefs and concerns.

The opportunities for misunderstanding increase when the circumstances of debate restrict candor. For example, some critics of nuclear power actually believe that the technology can be operated with reasonable safety. However, they oppose it because they believe that its costs and benefits are distributed inequitably. Although they might like to discuss these issues, critics find that public hearings about risk and safety often provide them with their only forum for venting their concern. If they oppose the technology, then they are forced to do so on safety grounds, even if this means misrepresenting their perceptions of the actual risk.

Individuals also have difficulty detecting inconsistencies in their own beliefs or realizing how simple reformulations would change their perspective on issues. For example, most people would prefer a gamble with a 25 percent chance of losing $200 (and a 75 percent chance of losing nothing) to a gamble with a sure loss of $50. Most of the same people would also buy a $50 insurance policy to protect against such a loss. What they will do depends on whether the $50 is described as a sure loss or as an insurance premium. As a result, one cannot predict how people will respond to an issue without knowing how they will perceive it, which depends, in turn, on how it will be presented to them by merchandisers, politicians, or the media.

Thus, people's insensitivity to the importance of how risk issues are presented exposes them to manipulation. For example, a risk might seem much worse when described in relative terms than in absolute terms (e.g., doubling their risk versus increasing that risk from 1 in a million to 1 in a half million). Although both representations of the risk might be honest, their impacts would be quite different. Perhaps the only fair approach is to present the risk from both perspectives, letting recipients determine which one (or which hybrid) best represents their world view.

SUMMARY

These statements (and others like them cited elsewhere in this appendix) reduce both complex people and intricate research literatures to necessarily oversimplified summaries. Neither the people nor the literature can be read without their appropriate context. Much of Section II discussed the intricacies of the literature and the sort of conclusions than might be extracted from it. Much of this whole appendix concerns the context for risk perception. Ideally, one would have polished studies of how specific people respond to specific risks, either in messages or in the flesh (or the metal). Those should be the standards for designing and evaluating risk communication programs. In lieu of such studies, such principles are all that we have to go on. They are the stuff of everyday explanations of behavior. They can be enriched, refined, and (sometimes) disqualified by behavioral research.

VII
CONCLUSION

INDIVIDUAL LEARNING

Making decisions about risks is often complex, whether done individually or as part of a larger social-political process. So is dealing with many of life's other decisions, even without obvious risks to health and safety (e.g., choosing a career, a partner, an anniversary present). All these decisions have sets of options to consider, bodies of fact to master, and competing objectives to weigh. Adding to the complexity of these individual decisions is the fact that each of us confronts so many of them—each with its own details and nuances.

Individually and collectively, these decisions present a daunting challenge to identify those courses of action that are in our own best interests. It should not be surprising if people sometimes feel overwhelmed by the panoply of risks thrown at them, sometimes seem to respond suboptimally, and sometimes get angry at those who force them to deal with yet another risk—even if it is associated with a technology bringing considerable benefit.

However, although the substance of these decisions may vary enormously, their common elements mean that there is an opportunity for learning some general lessons from this experience with diverse risks. So, even though few people receive formal training in decision-making methods, life itself can provide an education. People could not make it through life if they had not learned something about the relative riskiness of different activities (e.g., driving at night versus driving during the day, getting polio from vaccine versus getting it while unvaccinated, storing household chemicals under the sink versus storing them out of the reach of children). People would be perennially dissatisfied if they had not acquired some ability to understand and predict their own tastes. A representative democracy could not function if people did not have some ability to evaluate the candor and competence of political candidates and governmental officials. There would not be significant declines in smoking and fat consumption if people were not able to extract personally relevant implications from risk communications.

Some of these accomplishments are documented in the references cited in the preceding sections. Most are also common knowledge (although perhaps not as precisely delineated as they can be in systematic research). Most are also incomplete. Both anecdotal and systematic observations can point to places where people misestimate risks, mistake their own needs, misjudge public figures, or misinterpret the message of risk communications. In some cases, this is because life is not structured for learning. It may not provide people with prompt, immediate feedback on how well they are doing. It may discourage them from admitting the need to learn (without which even the sharpest feedback may have little value).

Under these circumstances, a guide like this can facilitate learning in several ways. One is to provide a structure for thinking about risk controversies, so as to facilitate identifying common elements and extracting general lessons. A second is to summarize the lessons found in the research literature and in the pooled experience of risk communicators (and communicants). In some cases, these lessons will confirm readers' expectations; in others, they will suggest alternative interpretations; in still others, they will raise issues that have not been considered. A third way is to provide annotated references to the research literature that could be consulted for more detailed treatment of specific risk issues. Making this research generally available in nontechnical terms can help to level the playing field, by granting equal access to it for all parties to risk controversies (and not just for those parties with staffs paid to follow the research literature).

Finally, such a guide can provide some insight into the psychological processes of the parties involved in risk controversies. That insight can be used directively, by those who must design risk communications and interpret the responses of the public to them. It can also be used reflectively, by those who wish to clarify the psychological limits to their own participation in risk management. These groups include nontechnical people concerned about interpreting the nature of risks, as well as technical people concerned about making themselves understood to others.

Such understanding has both a "cognitive" and a "motivational" component (to use psychological jargon for a moment). That is, it involves both how people think and how people feel. Deciphering scientific communications can be complicated both by difficulty interpreting strange terms or unfamiliar units (e.g., very small probabilities) and by difficulty coping with one's anger with the risk communicators (e.g., for their perceived insensitivity or vested interests). Designing such communications can be complicated both by difficulty interpreting complex social processes and by difficulty

managing one's frustration at being mistrusted and disbelieved. Better risk communication is typically thought of as a largely cognitive enterprise, focused on conveying factual material more comprehensibly. Accomplishing that goal requires an understanding of what aspects of risk conflicts really hinge on scientific facts. If it can be accomplished, then risk conflicts can be focused on areas of legitimate disagreement, without the confusion and frustration generated by the receipt of incomprehensible messages. Such messages both blur the issues and create the feeling that communicators care so little—or live in such a different world—that they cannot communicate in ways that address recipients' needs.

SOCIETAL LEARNING

Sweeping statements about people and society are easy to make, but hard to substantiate. If I were to chance a summary of personal observations from 15 years of working on this topic, it would be that there is increasing sophistication on the part of all concerned. We have better risk science than we had in the past and a better understanding of its limits. We have increasing understanding among risk managers of the need to take public concerns seriously when designing risk policies and among members of the public when deciding which risks to worry about and how to worry about them. We have increasing professionalism in reporting about risk issues and increasing ability to read or view risk stories with a discerning eye.

We also have, however, a long way to go in each of these respects. Moreover, the learning to date has come at a price that creates an obstacle to future progress. People remember their own past mistakes (at least the more obvious ones), which makes them hesitant about future actions. They also remember others' mistakes (at least those from which they think they have suffered), which makes them leery of those others' future actions. It is hard to erase a shadow of doubt or undo the undue impact of first impressions.

As in a social relationship, by the time those involved learn how to get along with a significant other, they may have hurt one another enough that they cannot apply these lessons in that relationship. Unfortunately, industry cannot break off its relationship with its current public (or its current government or current media) and start up with a new, more enlightened one. So, some personal wounds need to heal at the same time as we are collectively addressing new problems.

In addition, old problems continue to aggravate these wounds and to undermine the parties' faith in one another. For example, the question of whether to complete or operate many nuclear reactors is a lingering source of mutual frustration among all involved. The public commitments made by the various parties concerned are such that the conflicts have a life of their own. They may defy reasoned resolution and be almost refractory to the addition of scientific evidence. The strategizing and posturing of the parties may make great sense when viewed as part of a political struggle. Yet when viewed as part of a disciplined debate over risks and benefits, they can strengthen perceptions of a callous industry and hysterical public.

A guide such as this cannot dispel such complex conflicts and emotions. They are natural and legitimate parts of life. It can, however, help to put them in perspective, leaving the conflicts that remain better focused and more productive.

BIBLIOGRAPHY

Alfidi, J. 1971. Informed consent: A study of patient reaction. Journal of the American Medical Association 216:1325–1329.

Appelbaum, P. S., C. W. Lidz, and A. Meisel. 1987. Informed Consent: Legal Theory and Clinical Practice. New York: Oxford University Press.

Appelbaum, R. P. 1977. The future is made, not predicted: Technocratic planners vs. public interests. Society (May/June):49–53.

Applied Management Sciences. 1978. Survey of consumer perceptions of patient package inserts for oral contraceptives. NTIS No. PB-248-740. Washington, D.C.: Applied Management Sciences.

Armstrong, J. S. 1975. Tom Swift and his electric regression analysis machine 1973. Psychological Reports 36:806.

Atkinson, R. C., R. J. Herrnstein, G. Lindzey, and R. D. Luce. 1988. Stevens' Handbook of Experimental Psychology. New York: Wiley Interscience.

Bar-Hillel, M. 1973. On the subjective probability of compound events. Organizational Behavior and Human Performance 9:396–406.

Bar-Hillel, M. 1980. The base rate fallacy in probability judgment. Acta Psychologica 44:211–233.

Barber, W. C. 1979. Controversy plagues setting of environmental standards. Chemical and Engineering News 57(17):34–37.

Barraclough, G. 1972. Mandarins and Nazis. New York Review of Books 19(6):37–42.

Bazelon, D. L. 1979. Risk and responsibility. Science 205(4403):277–280.

Bentkover, J. D., V. T. Covello, and J. Mumpower, eds. 1985. Benefits Assessment: The State of the Art. Dordrecht, Holland: D. Reidel.

Berkson, J., T. B. Magath, and M. Hurn. 1939–1940. The error of estimate of the blood cell count as made with the hemocytometer. American Journal of Physiology 128:309–323.

Beyth-Marom, R. 1982. How probable is probable? Journal of Forecasting 1:257–269.

Bick, T., C. Hohenemser, and R. W. Kates. 1979. Target: Highway risks. Environment 21(2):7–15, 29–38.

Bickerstaffe, J., and D. Peace. 1980. Can there be a consensus on nuclear power? Social Studies of Science 10:309–344.

Bradburn, N. M., and S. Sudman. 1979. Improving Interview Method and Questionnaire Design. San Francisco: Jossey-Bass.

Brokensha, D. W., D. M. Warren, and O. Werner. 1980. Indigenous Knowledge: Systems and Development. Lanham, Md.: University Press of America.

Brookshire, D. S., B. C. Ives, and W. D. Schulze. 1976. The valuation of aesthetic preferences. Journal of Environmental Economics and Management 3:325–346.

Brown, R. 1965. Social Psychology. Glencoe, Ill.: Free Press.

Burton, I., R. W. Kates, and G. F. White. 1978. The Environment as Hazard. New York: Oxford University Press.

Callen, E. 1976. The science court. Science 193:950–951.

Campbell, D. T. 1975. Degrees of freedom and the case study. Comparative Political Studies 8:178–193.

Campbell, D. T., and A. Erlebacher. 1970. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In Compensatory Education: A National Debate, Vol. 3, Disadvantaged Child, J. Hellmuth, ed. New York: Brunner/Mazel.
Campen, J. 1985. Benefit-Cost and Beyond. Cambridge, Mass.: Ballinger.
Carterette, E. C., and M. P. Friedman. 1974. Handbook of Perception, Vol. 2. New York: Academic Press.
Chapman, L. J., and J. P. Chapman. 1969. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology 74:271–280.
Chemical and Engineering News. 1980. A look at human error. 58(18):82.
Cohen, B., and I. Lee. 1979. A catalog of risks. Health Physics 36:707–722.
Cohen, J. 1962. The statistical power of abnormal-social psychological research: A review. Journal of Abnormal and Social Psychology 65(3):145–153.
Commoner, B. 1979. The Politics of Energy. New York: Knopf.
Conn, W. D., ed. 1983. Energy and Material Resources. Boulder, Colo.: Westview.
Cotgrove, A. 1982. Catastrophe or Cornucopia? The Environment, Politics and the Future. New York: John Wiley & Sons.
Covello, V. T., P. M. Sandman, and P. Slovic. 1988. Risk Communication, Risk Statistics, and Risk Comparisons: A Manual for Plant Managers. Washington, D.C.: Chemical Manufacturers Association.
Covello, V., D. von Winterfeldt, and P. Slovic. 1986. Risk communication: A review of the literature. Risk Abstracts 3(4):171–182.
Crask, M. R., and W. D. Parreault, Jr. 1977. Validation of discriminant analysis in marketing research. Journal of Marketing Research 14:60–68.
Crouch, E. A. C., and R. Wilson. 1982. Risk/Benefit Analysis. Cambridge, Mass.: Ballinger.
Cummings, R. G., D. S. Brookshire, and W. D. Schulze, eds. 1986. Valuing Environmental Goods: An Assessment of the Contingent Valuation Method. Totowa, N.J.: Rowman & Allanheld.
Davidshofer, I. O. 1976. Risk-taking and vocational choice: Reevaluation. Journal of Counseling Psychology 23:151–154.
Davis, J. 1969. Group Performance. Reading, Mass.: Addison-Wesley.
Dietz, T. M., and R. W. Rycroft. 1987. The Risk Professionals. Washington, D.C.: Russell Sage Foundation.
Doern, G. B. 1978. Science and technology in the nuclear regulatory process: The case of Canadian uranium miners. Canadian Public Administration 21:51–82.
Dreman, D. 1979. Contrarian Investment Strategy. New York: Random House.
Driver, B., G. Peterson, and R. Gregory, eds. 1988. Evaluative Amenity Resources. New York: Venture.
Dunlap, T. R. 1978. Science as a guide in regulating technology: The case of DDT in the United States. Social Studies of Science 8:265–285.
Eiser, J. R., ed. 1982. Social Psychology and Behavioral Medicine. New York: John Wiley & Sons.
Elliot, G. R., and C. Eisdorfer. 1982. Stress and Human Health. New York: Springer-Verlag.

Fairley, W. B. 1977. Evaluating the "small" probability of a catastrophic accident from the marine transportation of liquefied natural gas. In Statistics and Public Policy, W. B. Fairley and F. Mosteller, eds. Reading, Mass.: Addison-Wesley.
Feller, W. 1968. An Introduction to Probability Theory and Its Applications, 3d ed., Vol. 1. New York: John Wiley & Sons.
Fineberg, H. V. 1988. Education to prevent AIDS: Prospects and obstacles. Science 239(4840):592–596.
Fischer, D. H. 1970. Historians' Fallacies. New York: Harper & Row.
Fischhoff, B. 1980. For those condemned to study the past: Reflections on historical judgment. In New Directions for Methodology of Behavior Science: Fallible Judgment in Behavioral Research, R. A. Shweder and D. W. Fiske, eds. San Francisco: Jossey-Bass.
Fischhoff, B. 1981. Informed consent for transient nuclear workers. In Equity Issues in Nuclear Waste Management, R. Kasperson and R. W. Kates, eds. Cambridge, Mass.: Oelgeschlager, Gunn and Hain.
Fischhoff, B. 1983. "Acceptable risk": The case of nuclear power. Journal of Policy Analysis and Management 2(4):559–575.
Fischhoff, B. 1984. Setting standards: A systematic approach to managing public health and safety risks. Management Science 30:823–843.
Fischhoff, B. 1985a. Managing risk perceptions. Issues in Science and Technology 2(1):83–96.
Fischhoff, B. 1985b. Protocols for environmental reporting: What to ask the experts. The Journalist (Winter):11–15.
Fischhoff, B. 1985c. Risk analysis demystified. NCAP News (Winter):30–33.
Fischhoff, B. 1987. Treating the public with risk communications: A public health perspective. Science, Technology, and Human Values 12:3–19.
Fischhoff, B. 1988. Judgment and decision making. In The Psychology of Human Thought, R. J. Sternberg and E. E. Smith, eds. New York: Cambridge University Press.
Fischhoff, B., and L. A. Cox, Jr. 1985. Conceptual framework for regulatory benefits assessment. In Benefits Assessment: The State of the Art, J. D. Bentkover, V. T. Covello, and J. Mumpower, eds. Dordrecht, Holland: D. Reidel.
Fischhoff, B., and L. Furby. 1988. Measuring values: A conceptual framework for interpretive transactions with special reference to contingent valuation of visibility. Journal of Risk and Uncertainty 1:147–184.
Fischhoff, B., and D. MacGregor. 1983. Judged lethality: How much people seem to know depends upon how they are asked. Risk Analysis 3:229–236.
Fischhoff, B., and O. Svenson. 1987. Perceived risks of radionuclides: Understanding public understanding. In Radionuclides in the Food Chain, G. Schmidt, ed. New York: Praeger.
Fischhoff, B., L. Furby, and R. Gregory. 1987. Evaluating voluntary risks of injury. Accident Analysis and Prevention 19(1):51–62.
Fischhoff, B. S., Lichtenstein, P. Slovic, S. L. Derby, and R. L. Keeney. 1981. Acceptable Risk. New York: Cambridge University Press.
Fischhoff, B., P. Slovic, and S. Lichtenstein. 1978. Fault trees: Sensitivity of assessed failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance 4:330–344.

Fischhoff, B., P. Slovic, and S. Lichtenstein. 1980. Knowing what you want: Measuring labile values. In Cognitive Processes in Choice and Decision Behavior, T. Wallsten, ed. Hillsdale, N.J.: Erlbaum.
Fischhoff, B., P. Slovic, and S. Lichtenstein. 1981. Lay foibles and expert fables in judgments about risk. In Progress in Resource Management and Environmental Planning, T. O'Riordan and R. K. Turner, eds. New York: John Wiley & Sons.
Fischhoff, B., P. Slovic, S. Lichtenstein, S. Read, and B. Combs. 1978. How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. Policy Sciences 9:127–152.
Fischhoff, B., S. R. Watson, and C. Hope. 1984. Defining risk. Policy Sciences 17:123–129.
Fiske, S., and S. Taylor. 1984. Social Cognition. Reading, Mass.: Addison-Wesley.
Frankel, C. 1974. The rights of nature. In When Values Conflict, C. Schelling, J. Voss, and L. Tribe, eds. Cambridge, Mass.: Ballinger.
Friedman, S. M. 1981. Blueprint for breakdown: Three Mile Island and the media before the accident. Journal of Communication 31:116–129.
Furby, L., and B. Fischhoff. In press. Rape self-defense strategies: A review of their effectiveness. Victimology.
Gamble, D. J. 1978. The Berger Inquiry: An impact assessment process. Science 199(4332):946–951.
Gilovich, T., R. Vallone, and A. Tversky. 1985. The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology 17:295–314.
Gotchy, R. L. 1983. Health risks from the nuclear fuel cycle. In Health Risks of Energy Technologies, C. C. Travis and E. L. Etnier, eds. Boulder, Colo.: Westview.
Green, A. E., and A. J. Bourne. 1972. Reliability Technology. New York: Wiley Interscience.
Hackney, J. D., and W. S. Linn. 1984. Human toxicology and risk assessment. In Handbook on Risk Assessment. Washington, D.C.: National Science Foundation.
Hammond, K. R., and L. Adelman. 1976. Science, values and human judgment. Science 194:389–396.
Hance, B. J., C. Chess, and P. M. Sandman. 1988. Improving Dialogue with Communities: A Risk Communication Manual for Government. Trenton: Division of Science and Research Risk Communication Unit, New Jersey Department of Environmental Protection.
Handler, P. 1980. Public doubts about science. Science 208(4448):1093.
Hanley, J. 1980. The silence of scientists. Chemical and Engineering News 58(12):5.
Harris, L. 1980. Risk in a complex society. Public opinion survey conducted for Marsh and McLennan Companies, Inc.
Harriss, R., and C. Hohenemser. 1978. Mercury: Measuring and managing risk. Environment 20(9).
Hasher, L., and R. T. Zacks. 1984. Automatic and effortful processes in memory. Journal of Experimental Psychology: General 108:356–388.
Henrion, M., and B. Fischhoff. 1986. Assessing uncertainty in physical constants. American Journal of Physics 54(9):791–798.
Henshel, R. L. 1975. Effects of disciplinary prestige on predictive accuracy: Distortions from feedback loops. Futures 7:92–196.

Herbert, J. H., L. Swanson, and P. Reddy. 1979. A risky business. Environment 21(6):28–33.
Hershey, J. C., and P. J. H. Schoemaker. 1980. Risk taking and problem context in the domain of losses: An expected utility analysis. Journal of Risk and Insurance 47:111–132.
Hirokawa, R. Y., and M. S. Poole. 1986. Communication and Group Decison Making. Beverly Hills, Calif.: Sage.
Hohenemser, K. H. 1975. The failsafe risk. Environment 17(1):6–10.
Holden, C. 1980. Love Canal residents under stress. Science 208:1242–1244.
Hovland, C. I., I. L. Janis, and H. H. Kelley. 1953. Communication and Persuasion: Psychological Studies of Opinion Change. New Haven, Conn.: Yale University Press.
Hynes, M., and E. Vanmarcke. 1976. Reliability of embankment performance prediction. In Proceedings of the ASCE Engineering Mechanics Division Specialty Conference. Waterloo, Ontario, Canada: University of Waterloo Press.
Ingram, M. J., D. J. Underhill, and T. M. L. Wigley. 1978. Historical climatology. Nature 276:329–334.
Inhaber, H. 1979. Risk with energy from conventional and nonconventional sources. Science 203(4382):718–723.
Institute of Medicine. 1986. Confronting AIDS: Directions for Public Health, Health Care, and Research. Washington, D.C.: National Academy Press.
James, W. 1988. Baseball Abstract. New York: Ballantine.
Janis, I. L., ed. 1982. Counseling on Personal Decisions. New Haven, Conn.: Yale University Press.
Jennergren, L. P., and R. L. Keeney. 1982. Risk assessment. In Handbook of Applied Systems Analysis. Laxenburg, Austria: International Institute of Applied Systems Analysis.
Johnson, B. B., and V. T. Covello, eds. 1987. The Social and Cultural Construction of Risk: Essays on Risk Selection and Perception. Dordrecht, Holland: D. Reidel.
Joksimovich, V. 1984. Models in risk assessment for hazard characterization. In Handbook of Risk Assessment. Washington, D.C.: National Science Foundation.
Joubert, P., and L. Lasagna. 1975. Commentary: Patient package inserts. Clinical Pharmacology and Therapeutics 18(5):507–513.
Kadlec, R. 1984. Field and laboratory event investigation for hazard characterization. In Handbook of Risk Assessment. Washington, D.C.: National Science Foundation.
Kahneman, D., and A. Tversky. 1972. Subjective probability: A judgment of representativeness. Cognitive Psychology 3:430–454.
Kasperson, R. 1986. Six propositions on public participation and their relevance for risk communication. Risk Analysis 6(3):275–281.
Keeney, R. L. 1980. Siting Energy Facilities. New York: Academic Press.
Keeney, R. L., and H. Raiffa. 1976. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. New York: John Wiley & Sons.
Kolata, G. B. 1980. Love Canal: False alarm caused by botched study. Science 208(4449):1239–1242.
Koriat, A., S. Lichtenstein, and B. Fischhoff. 1980. Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory 6:107–118.

Krohn, W., and P. Weingart. 1987. Commentary: Nuclear power as a social experiment—European political "fall-out" from the Chernobyl meltdown. Science, Technology, and Human Values 12(2):52–58.

Kunce, J. T., D. W. Cook, and D. E. Miller. 1975. Random variables and correlational overkill. Educational and Psychological Measurement 35:529–534.

Kunreuther, H., R. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, and N. Katz. 1978. Disaster Insurance Protection. New York: John Wiley & Sons.

Lachman, R., J. T. Lachman, and E. C. Butterfield. 1979. Cognitive Psychology and Information Processing. Hillsdale, N.J.: Erlbaum.

Lakatos, I. 1970. Falsification and scientific research programmes. In Criticism and the Growth of Scientific Knowledge, I. Lakatos and A. Musgrave, eds. New York: Cambridge University Press.

Lanir, Z. 1982. Strategic Surprises. Tel Aviv, Israel: Hakibbutz Hameuchad.

Lave, L. B. 1978. Ambiguity and inconsistency in attitudes toward risk: A simple model. Pp. 108–114 in Proceedings of the Society for General Systems Research Annual Meeting. Louisville, Ky.: Society for General Systems Research.

Lawless, E. W. 1977. Technology and Social Shock. New Brunswick, N.J.: Rutgers University Press.

Lazarsfeld, P. 1949. The American soldier—An expository review. Public Opinion Quarterly 13:377–404.

Levine, M. 1974. Scientific method and the adversary model: Some preliminary thoughts. American Psychologist 29:661–716.

Lichtenstein, S., and B. Fischhoff. 1980. Training for calibration. Organizational Behavior and Human Performance 26:149–171.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips. 1982. Calibration of probabilities: The state of the art. In Judgment Under Uncertainty: Heuristics and Biases, P. Slovic and A. Tversky, eds. New York: Cambridge University Press.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs. 1978. Judged frequency of lethal events. Journal of Experimental Psychology: Human Learning and Memory 4:551–578.

Lindman, H. G., and W. Edwards. 1961. Supplementary report: Unlearning the gambler's fallacy. Journal of Experimental Psychology 62:630.

Linville, P., B. Fischhoff, and G. Fischer. 1988. Judgments of AIDS Risks. Pittsburgh, Pa.: Carnegie-Mellon University, Department of Social and Decision Sciences.

MacLean, D. 1987. Understanding the nuclear power controversy. In Scientific Controversies: Case Studies in the Resolution and Closure of Disputes in Science and Technology, H. T. Engelhardt, Jr., and A. L. Caplan, eds. New York: Cambridge University Press.

Markovic, M. 1970. Social determinism and freedom. In Mind, Science and History, H. E. Keifer and M. K. Munitz, eds. Albany: State University of New York Press.

Martin, E. 1980. Surveys as Social Indicators: Problems in Monitoring Trends. Chapel Hill: Institute for Research in Social Science, University of North Carolina.

Mazur, A. 1973. Disputes between experts. Minerva 11:243–262.

Mazur, A. 1981. The Dynamics of Technical Controversy. Washington, D.C.: Communications Press.

Mazur, A., A. A. Marino, and R. O. Becker. 1979. Separating factual disputes from value disputes in controversies over technology. Technology in Society 1:229–237.

McGrath, P. E. 1974. Radioactive Waste Management: Potentials and Hazards From a Risk Point of View. Report EUR FNR-1204 (KFK 1992). Karlsruhe, West Germany: US-EURATOM Fast Reactor Program.

McNeil, B. J., R. Weichselbaum, and S. G. Pauker. 1978. The fallacy of the 5-year survival rate in lung cancer. New England Journal of Medicine 299:1397–1401.

Morgan, M. 1986. Conflict and confusion: What rape prevention experts are telling women. Sexual Coercion and Assault 1(5):160–168.

Murphy, A. H., and B. G. Brown. 1983. Forecast terminology: Composition and interpretation of public weather forecasts. Bulletin of the American Meteorological Society 64:13–22.

Murphy, A. H., and R. L. Winkler. 1984. Probability of precipitation forecasts. Journal of the American Statistical Association 79:391–400.

National Research Council. 1976. Surveying Crime. Washington, D.C.: National Academy Press.

National Research Council. 1982. Survey Measure of Subjective Phenomena. Washington, D.C.: National Academy Press.

National Research Council. 1983a. Priority Mechanisms for Toxic Chemicals. Washington, D.C.: National Academy Press.

National Research Council. 1983b. Risk Assessment in the Federal Government: Managing the Process. Washington, D.C.: National Academy Press.

Nelkin, D. 1977. Technological Decisions and Democracy. Beverly Hills, Calif.: Sage.

Nelkin, D., ed. 1984. Controversy: Politics of Technical Decisions. Beverly Hills, Calif.: Sage.

Neyman, J. 1979. Probability models in medicine and biology: Avenues for their validation for humans in real life. Berkeley: University of California, Statistical Laboratory.

Nisbett, R. E., and L. Ross. 1980. Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, N.J.: Prentice-Hall.

Northwest Coalition for Alternatives to Pesticides. 1985. Position Document—Risk Analysis. NCAP News (Winter):33.

Office of Science and Technology Policy. 1984. Chemical carcinogens: Review of the science and its associated principles. Federal Register 49(100):21594–21661.

O'Flaherty, E. J. 1984. Pharmacokinetic methods in risk assessment. In Handbook of Risk Assessment. Washington, D.C.: National Science Foundation.

O'Leary, M. K., W. D. Coplin, H. B. Shapiro, and D. Dean. 1974. The quest for relevance. International Studies Quarterly 18:211–237.

Östberg, G., H. Hoffstedt, G. Holm, B. Klingernstierna, B. Rydnert, V. Samsonowitz, and L. Sjöberg. 1977. Inconceivable Events in Handling Material in Heavy Mechanical Engineering Industry. Stockholm, Sweden: National Defense Research Institute.

Otway, H. J., and D. von Winterfeldt. 1982. Beyond acceptable risk: On the social acceptability of technologies. Policy Sciences 14:247–256.

Page, T. 1978. A generic view of toxic chemicals and similar risks. Ecology Law Quarterly 7:207–243.
Page, T. 1981. A framework for unreasonable risk in the Toxic Substances Control Act. In Carcinogenic Risk Assessment, R. Nicholson, ed. New York: New York Academy of Sciences.
Parducci, A. 1974. Contextual effects: A range-frequency analysis. In Handbook of Perception, Vol. 2, E. C. Carterette and M. P. Friedman, eds. New York: Academic Press.
Payne, S. L. 1952. The Art of Asking Questions. Princeton, N.J.: Princeton University Press.
Pearce, D. W. 1979. Social cost-benefit analysis and nuclear futures. In Energy Risk Management, G. T. Goodman and W. D. Rowe, eds. New York: Academic Press.
Peterson, C. R., and L. R. Beach. 1967. Man as an intuitive statistician. Psychological Bulletin 69(1):29–46.
Peto, R. 1980. Distorting the epidemiology of cancer. Nature 284:297–300.
Pew, R. D., C. Miller, and C. E. Feeher. 1982. Evaluation of Proposed Control Room Improvements Through Analysis of Critical Operator Decisions. Palo Alto, Calif.: Electric Power Research Institute.
Pinder, G. F. 1984. Groundwater contaminant transport modeling. Environmental Science and Technology 18(4):108A–114A.
Poulton, E. C. 1968. The new psychophysics: Six models of magnitude estimation. Psychological Bulletin 69:1–19.
Poulton, E. C. 1977. Quantitative subjective assessments are almost always biased, sometimes completely misleading. British Journal of Psychology 68:409–421.
President's Commission on the Accident at Three Mile Island. 1979. Report of the President's Commission on the Accident at Three Mile Island. Washington, D.C.: U.S. Government Printing Office.
Rayner, S., and R. Cantor. 1987. How fair is safe enough?: The cultural approach to societal technology choice. Risk Analysis 7(1):3–9.
Reissland, J., and V. Harries. 1979. A scale for measuring risks. New Scientist 83:809–811.
Rodricks, J. V., and R. G. Tardiff. 1984. Animal research methods for dose-response assessment. In Handbook of Risk Assessment. Washington, D.C.: National Science Foundation.
Rokeach, M. 1973. The Nature of Human Values. New York: The Free Press.
Roling, G. T., L. W. Pressgrove, E. B. Keefe, and S. B. Raffin. 1977. An appraisal of patients' reactions to "informed consent" for peroral endoscopy. Gastrointestinal Endoscopy 24(2):69–70.
Rosencranz, A., and G. S. Wetstone. 1980. Acid precipitation: National and international responses. Environment 22(5):6–20, 40–41.
Rosenthal, R., and R. L. Rosnow. 1969. Artifact in Behavioral Research. New York: Academic Press.
Rothman, S., and S. R. Lichter. 1987. Elite ideology and risk perception in nuclear energy policy. American Political Science Review 81(2):383–404.
Rothschild, N. M. 1978. Rothschild: An antidote to panic. Nature 276:555.
Rubin, D., and D. Sachs, eds. 1973. Mass Media and the Public. New York: Praeger.
Schnaiburg, A. 1980. The Environment: From Surplus to Scarcity. New York: Oxford University Press.

Schneider, S. H., and L. E. Mesirow. 1976. The Genesis Strategy. New York: Plenum.
Schneiderman, M. A. 1980. The uncertain risks we run: Hazardous material. In Societal Risk Assessment: How Safe is Safe Enough?, R. C. Schwing and W. A. Albers, Jr., eds. New York: Plenum.
Schudson, M. 1978. Discovering the News. New York: Basic Books.
Schwarz, E. D. 1978. The use of a checklist in obtaining informed consent for treatment with medicare. Hospital and Community Psychiatry 29:97–100.
Seligman, M. E. P. 1975. Helplessness. San Francisco: Freeman, Cooper.
Shaklee, H., B. Fischhoff, and L. Furby. 1988. The psychology of contraceptive surprises: Cumulative risk and contraceptive failure. Eugene, Oreg.: Eugene Research Institute.
Sharlin, H. I. 1987. Macro-risks, micro-risks, and the media: The EDB case. In The Social and Cultural Construction of Risk, B. B. Johnson and V. T. Covello, eds. Dordrecht, Holland: D. Reidel.
Sheridan, T. B. 1980. Human error in nuclear power plants. Technology Review 82(4):23–33.
Shroyer, T. 1970. Toward a critical theory for advanced industrial society. In Recent Sociology, Vol. 2, Patterns of Communicative Behavior, H. P. Drietzel, ed. London: Macmillan.
Sioshansi, F. P. 1983. Subjective evaluation using expert judgment: An application. IEEE Transactions on Systems, Man and Cybernetics 13(3):391–397.
Sjöberg, L. 1979. Strength of belief and risk. Policy Sciences 11:539–573.
Slovic, P. 1962. Convergent validation of risk-taking measures. Journal of Abnormal and Social Psychology 65:68–71.
Slovic, P. 1986. Informing and educating the public about risk. Risk Analysis 6(4):403–415.
Slovic, P., and B. Fischhoff. 1977. On the psychology of experimental surprises. Journal of Experimental Psychology: Human Perception and Performance 3:544–551.
Slovic, P., and B. Fischhoff. 1983. How safe is safe enough? Determinants of perceived and acceptable risk. In Too Hot to Handle? Social and Policy Issues in the Management of Radioactive Wastes, C. Walker, L. Gould, and E. Woodhouse, eds. New Haven, Conn.: Yale University Press.
Slovic, P., B. Fischhoff, and S. Lichtenstein. 1978. Accident probabilities and seatbelt usage: A psychological perspective. Accident Analysis and Prevention 17:10–19.
Slovic, P., B. Fischhoff, and S. Lichtenstein. 1979. Rating the risks. Environment 21:14–20, 30, 36–39.
Slovic, P., B. Fischhoff, and S. Lichtenstein. 1980. Facts vs. fears: Understanding perceived risk. In Societal Risk Assessment: How Safe Is Safe Enough?, R. Schwing and W. A. Albers, Jr., eds. New York: Plenum.
Slovic, P., B. Fischhoff, and S. Lichtenstein. 1984. Modeling the societal impact of fatal accidents. Management Science 30:464–474.
Slovic, P., B. Fischhoff, S. Lichtenstein, B. Corrigan, and B. Combs. 1977. Preference for insuring against probable small losses: Implications for the theory and practice of insurance. Journal of Risk and Insurance 44:237–258.
Smith, V. K., and W. H. Desvousges. 1986. Measuring Water Quality Benefits. Boston: Kluwer.
Stallen, P. J. 1980. Risk of science or science of risk? In Society, Technology and Risk Assessment, J. Conrad, ed. London: Academic Press.

Starr, C. 1969. Social benefit versus technological risk. Science 165:1232–1238.
Svenson, O. 1981. Are we all less risky and more skillful than our fellow drivers? Acta Psychologica 47:143–148.
Svenson, O., and B. Fischhoff. 1985. Levels of environmental decisions. Journal of Environmental Psychology 5:55–67.
Thompson, M. 1980. Aesthetics of risk: Culture or context. In Societal Risk Assessment, R. C. Schwing and W. A. Albers, Jr., eds. New York: Plenum.
Tockman, M. S., and A. M. Lilienfeld. 1984. Epidemiological methods in risk assessment. In Handbook of Risk Assessment. Washington, D.C.: National Science Foundation.
Travis, C. C. 1984. Modeling methods for exposure assessment. In Handbook of Risk Assessment. Washington, D.C.: National Science Foundation.
Tribe, L. H. 1972. Policy science: Analysis or ideology? Philosophy and Public Affairs 2:66–110.
Tukey, J. W. 1977. Some thoughts on clinical trials, especially problems of multiplicity. Science 198:679–690.
Tulving, E. 1972. Episodic and semantic memory. In Organization of Memory, E. Tulving and W. Donaldson, eds. New York: Academic Press.
Turner, C. F. 1980. Surveys of subjective phenomena. In The Measurement of Subjective Phenomena, D. Johnston, ed. Washington, D.C.: U.S. Government Printing Office.
Turner, C. F., and E. Martin, eds. 1985. Surveying Subjective Phenomena, Vols. 1 and 2. New York: Russell Sage Foundation.
Tversky, A., and D. Kahneman. 1971. The belief in the "law of small numbers." Psychological Bulletin 76:105–110.
Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. Cognitive Psychology 5:207–232.
Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. Science 185:1124–1131.
Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. Science 211(4481):453–458.
U.S. Committee on Government Operations. 1978. Teton Dam Disaster. Washington, D.C.: Government Printing Office.
U.S. Government. 1975. Hearings, 94th Cong., 1st Sess. Browns Ferry Nuclear Plant Fire, September 16, 1975. Washington, D.C.: U.S. Government Printing Office.
U.S. Nuclear Regulatory Commission. 1975. Reactor safety study: An assessment of accident risks in U.S. commercial nuclear power plants. WASH 1400 (NUREG-75/014). Washington, D.C.: U.S. Nuclear Regulatory Commission.
U.S. Nuclear Regulatory Commission. 1978. Risk Assessment Review Group to the U.S. Nuclear Regulatory Commission. NUREG/CR-0400. Washington, D.C.: U.S. Nuclear Regulatory Commission.
U.S. Nuclear Regulatory Commission. 1982. Safety Goals for Nuclear Power Plants: A Discussion Paper. NUREG-0880. Washington, D.C.: U.S. Nuclear Regulatory Commission.
U.S. Nuclear Regulatory Commission. 1983. PRA Procedures Guide. NUREG/CR-2300. Washington, D.C.: U.S. Nuclear Regulatory Commission.
Vlek, C. A. J., and P. J. Stallen. 1980. Rational and personal aspects of risk. Acta Psychologica 45:273–300.

Vlek, C. A. J., and P. J. Stallen. 1981. Judging risks and benefits in the small and in the large. Organizational Behavior and Human Performance 28:235–271.
von Winterfeldt, D., R. S. John, and K. Borcherding. 1981. Cognitive components of risk ratings. Risk Analysis 1(4):277–287.
Weaver, S. 1979. The passionate risk debate. The Oregon Journal, April 24.
Weinberg, A. M. 1979. Salvaging the atomic age. The Wilson Quarterly (Summer):88–112.
Weinstein, N. D. 1980a. Seeking reassuring or threatening information about environmental cancer. Journal of Behavioral Medicine 2:125–139.
Weinstein, N. D. 1980b. Unrealistic optimism about future life events. Journal of Personality and Social Psychology 93:806–820.
Weinstein, N. D., ed. 1987. Taking Care. New York: Cambridge University Press.
White, G., ed. 1974. Natural Hazards: Local, National and Global. New York: Oxford University Press.
Wilson, R. 1979. Analyzing the daily risks of life. Technology Review 81(4):40–46.
Wilson, V. L. 1980. Estimating changes in accident statistics due to reporting requirement changes. Journal of Safety Research 12(1):36–42.
Wohlstetter, R. 1962. Pearl Harbor: Warning and Decision. Stanford, Calif.: Stanford University Press.
Woodworth, R. S., and H. Schlosberg. 1954. Experimental Psychology. New York: Henry Holt.
Wortman, P. M. 1975. Evaluation research: A psychological perspective. American Psychologist 30:562–575.
Wynne, B. 1980. Technology, risk and participation. In Society, Technology and Risk Assessment, J. Conrad, ed. London: Academic Press.
Wynne, B. 1983. Institutional mythologies and dual societies in the management of risk. In The Risk Analysis Controversy, H. C. Kunreuther and E. V. Ley, eds. New York: Springer-Verlag.
Zeisel, H. 1980. Lawmaking and public opinion research: The President and Patrick Caddell. American Bar Foundation Research Journal 1:133–139.
Zentner, R. D. 1979. Hazards in the chemical industry. Chemical and Engineering News 57(45):25–27, 30–34.