Annu. Rev. Publ. Health 1993. 14:183–203

RISK PERCEPTION AND COMMUNICATION

Baruch Fischhoff
Department of Engineering and Public Policy, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Ann Bostrom
School of Public Policy, Georgia Institute of Technology, Atlanta, Georgia 30332-0345

Marilyn Jacobs Quadrel[1]
Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

KEY WORDS: health behavior, judgment, decision making

INTRODUCTION

Role of Risk Perceptions in Public Health

Many health risks are the result of deliberate decisions by individuals consciously trying to get the best deal possible for themselves and for those important to them. Some of these choices are private ones, such as whether to wear bicycle helmets and seatbelts, whether to read and follow safety warnings, whether to buy and use condoms, and how to select and cook food. Other choices involve societal issues, such as whether to protest the siting of hazardous waste incinerators and half-way houses, whether to vote for fluoridation and "green" candidates, and whether to support sex education in the schools.

In some cases, single choices can have a large effect on health risks (e.g. buying a car with airbags, taking a dangerous job, getting pregnant). In other

[1]Dr. Quadrel's current address is Decision Analysis Group, Battelle Pacific Northwest Laboratories, Richland, Washington 99352.

0163-7525/93/0510-0183$02.00

cases, the effects of individual choices are small, but can accumulate over multiple decisions (e.g. repeatedly ordering broccoli, wearing a seatbelt, using the escort service in parking garages). In still other cases, choices intended to affect health risks do nothing at all or the opposite of what is expected (e.g. responses to baseless cancer scares, adoption of quack treatments).

To make such decisions wisely, individuals need to understand the risks and the benefits associated with alternative courses of action. They also need to understand the limits to their own knowledge and the limits to the advice proffered by various experts. In this chapter, we review the research base for systematically describing a person's degree of understanding about health risk issues. We also consider some fundamental topics in designing and evaluating messages that are intended to improve that understanding. Following convention, we call these pursuits risk perception and risk communication research, respectively. In practice, the beliefs and messages being studied might deal with the benefits accompanying a risk, with the individuals and institutions who manage it, or with the broader issues that it raises (e.g. who gets to decide, how equitably risks and benefits are distributed).

The Role of Perceptions about Risk Perceptions in Public Health

The fundamental assumption of this chapter is that statements about other people's understanding must be disciplined by systematic data. People can be hurt by inaccuracies in their risk perceptions. They can also be hurt by inaccuracies in what various risk managers believe about those perceptions. Those managers might include physicians, nurses, public health officials, legislators, regulators, and engineers—all of whom have some say in what risks are created, what is communicated about them, and what role laypeople have in determining their fate.

If their understanding is overestimated, then people may be thrust into situations that they are ill-prepared to handle. If their understanding is underestimated, then people may be disenfranchised from decisions that they could and should make. The price of such misperceptions of risk perceptions may be exacted over the long run, as well as in individual decisions. The outcomes of health risk decisions partly determine people's physical and financial resources. The processes of health risk decisions partly determine people's degree of autonomy in managing their own affairs and in shaping their society.

In addition to citing relevant research results, the chapter emphasizes research methods. One conventional reason for doing so is improving access to material that is scattered over specialist literatures or part of the implicit knowledge conveyed in professional training. A second conventional reason

is to help readers evaluate the substantive results reported here, by giving a feeling for how they were produced.

A less conventional reason is to make the point that method matters. We are routinely struck by the strong statements made about other people's competence to manage risks, solely on the basis of anecdotal observation. These statements appear directly in pronouncements about, say, why people mistrust various technologies or fail to "eat right." Such claims appear more subtly in the myriad of health advisories, advertisements, and warnings directed at the public without any systematic evaluation. These practices assume that the communicator knows what people currently know, what they need to learn, what they want to hear, and how they will interpret a message.

Even the casual testing of a focus group shows a willingness to have those (smug) assumptions challenged.[1] The research methods presented here show the details needing attention and, conversely, the pitfalls to casual observation. The presentation also shows the limits to such research, in terms of how far current methods can go and how quickly they can get there. In our experience, once the case has been made for conducting behavioral research, it is expected to produce results immediately. That is, of course, a prescription for failure, and for undermining the perceived value of future behavioral research.

Overview

ORGANIZATION The following section, Quantitative Assessment, treats the most obvious question about laypeople's risk perceptions: Do they understand how big risks are? It begins with representative results regarding the quality of these judgments, along with some psychological theory regarding reasons for error. It continues with issues in survey design, which focus on how design choices can affect respondents' apparent competence. Some of these methodological issues reveal substantive aspects of lay risk perceptions.

The next section, Qualitative Assessment, shifts the focus from summary judgments to qualitative features of the events to which they are attached. It begins with the barriers to communication created when experts and laypeople unwittingly use terms differently. For example, when experts tell (or ask) people about the risks of drinking and driving, what do people think is meant regarding the kinds and amounts of "drinking" and of "driving"? The section continues by asking how people believe that risks "work," on the basis of which they might generate or evaluate control options.

The next section provides a general process for developing communications

[1]Focus groups are a popular technique in market research. In them, survey questions, commercial messages, or consumer products are discussed by groups of laypeople. Although they can generate unanticipated alternative interpretations, focus groups create a very different situation than that faced by an individual trying to make sense out of a question, message, or product (44).

about health risks. That process begins with identifying the information to be communicated, based on the descriptive study of what recipients know already and the formal analysis of what they need to know to make informed decisions. The process continues by selecting an appropriate format for presenting that information. It concludes with explicit evaluation of the resulting communication (followed by iteration if the results are wanting). The process is illustrated with examples taken from several case studies, looking at such diverse health risks as those posed by radon, Lyme disease, electromagnetic fields, carotid endarterechtomy, and nuclear energy sources in space.

EXCLUSIONS We do not address several issues that belong in a full account of their own, including the roles of emotion, individual differences (personality), culture, and social processes in decisions about risk. This set of restrictions suits the chapter's focus on how individuals think about risks. It may also suit a public health perspective, where it is often necessary to "treat" populations (with information) in fairly uniform ways. Access to these missing topics might begin with Refs. 27, 32, 36, 49, 66, 68, 71, 72.

QUANTITATIVE ASSESSMENT

Estimating the Size of Risks

A common presenting symptom in experts' complaints about lay decision making is that "laypeople simply do not realize how small (or large) the risk is." If that were the case, then the mission of risk communication would be conceptually simple (if technically challenging): Transmit credible estimates of how large the risks are (32, 49, 60, 68). Research suggests that lay estimates of risk are, indeed, subject to biases. Rather less evidence clearly implicates these biases in inappropriate risk decisions, or substantiates the idealized notion of people waiting for crisp risk estimates so that they can run well-articulated decision-making models. Such estimates are necessary, but not sufficient, for effective decisions.

In one early attempt to evaluate lay estimates of the size of risks, Lichtenstein et al (40) asked people to estimate the number of deaths in the US from 30 causes (e.g. botulism, tornados, motor vehicle accidents).[2] They

[2]The "people" in this study were members of the League of Women Voters and their spouses. Generally speaking, the people in the studies described here have been students paid for participation (hence, typically older than the proverbial college sophomores of some psychological research) or convenience samples of adults recruited through diverse civic groups (e.g. garden clubs, PTAs, bowling leagues). These groups have been found to differ more in what they think than in how they think. That is, their respective experiences have created larger differences in specific beliefs than in thought processes. Fuller treatment of sampling issues must await another opportunity.

used two different response modes, thus allowing them to check for the consistency of responses. One task presented pairs of causes; subjects chose the more frequent and then estimated the ratio of frequencies. The second task asked subjects to estimate the number of deaths in an average year; subjects were told the answer for one cause, in order to give an order-of-magnitude feeling (for those without a good idea for how many people live or die in the US in an average year). The study reached several conclusions that have been borne out by subsequent studies:

INTERNAL CONSISTENCY Estimates of relative frequency were quite consistent across response mode. Thus, people seemed to have a moderately well-articulated internal risk scale, which they could express even in unfamiliar response modes.

ANCHORING BIAS Direct estimates were influenced by the anchor given. Subjects told that 50,000 people die from auto accidents produced estimates two to five times higher than those produced by subjects told that 1000 die from electrocution. Thus, people seem to have less of a feel for absolute frequency, rendering them sensitive to the implicit cues in how questions are asked (51).

COMPRESSION Subjects' estimates showed less dispersion than did the statistical estimates. In this case, the result was an overestimation of small frequencies and an underestimation of large ones. However, the anchoring bias suggests that this pattern might have changed with different procedures, which would make the compression of estimates the more fundamental result.

AVAILABILITY BIAS At any level of statistical frequency, some causes of death consistently received higher estimates than others. These proved to be causes that are disproportionately visible (e.g. as reported in the news media, as experienced in subjects' lives). This bias seemed to reflect a general tendency to estimate the frequency of events by the ease with which they are remembered or imagined—while failing to realize what a fallible index such availability is (32, 65).

MISCALIBRATION OF CONFIDENCE JUDGMENTS In a subsequent study (21), subjects were asked how confident they were in their ability to choose the more frequent of the paired causes of death. They tended to be overconfident. For example, they had chosen correctly only 75% of the time when they were 90% confident of having done so. This result is a special case of a general tendency to be inadequately sensitive to the extent of one's knowledge (38, 72).

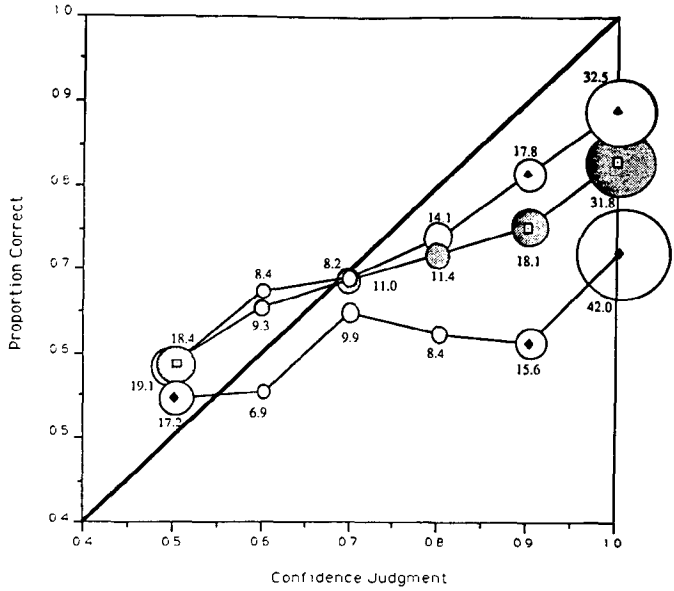

Figure 1 Calibration curves for adults (top, white: $N = 45$), not-at-risk teens (middle, dark: $N = 43$), and at-risk teens (bottom, white: $N = 45$). Each point indicates the proportion of correct answers among those in which subjects expressed a particular confidence level; the size of each circle indicates the percentage of answers held with that degree of confidence. (From Ref. 52.)

Figure 1 shows typical results from such a calibration test. In this case, subjects expressed their confidence in having chosen the correct answer to two-alternative questions regarding health behaviors [e.g. alcohol is (a) a depressant; (b) a stimulant]. The two curves reflect a group of middle-class adults and some of their adolescent children, recruited through school organizations.[3]

Response Mode Problems

One recurrent obstacle to assessing or improving laypeople's estimates of risk is reliance on verbal quantifiers. It is hard for them to know what experts

[3]In other studies comparing individuals drawn from these groups (53), we have also observed little difference in their respective response patterns. These studies suggest that any differences in their risk behaviors cannot be attributed to differences in the sorts of judgments considered in this chapter. If that is the case, and if such adults and teens do differ in their risk behaviors, then it may reflect differences in the benefits that they get from the behaviors (or in the risks and benefits of alternative behaviors).

mean when a risk is described as "very likely" or "rare"—or for experts to evaluate lay perceptions expressed in those terms. Such terms mean different things to different people, and even to the same person in different contexts (e.g. likely to be fatal versus likely to rain, rare disease versus rare Cubs pennant), sometimes even within communities of experts (3, 39, 67).

The Lichtenstein et al study (40) could observe the patterns reported above because it used an absolute response scale. As noted, it provided anchors to give subjects a feeling for how to answer. Doing so improved performance by drawing responses to the correct range, within which subjects were drawn to higher or lower values depending on the size of the anchor. Although most conclusions were relatively insensitive to these effects, they left no clear answer to the critical question of whether people overestimate or underestimate the risks that they face.

PERCEIVED LETHALITY A study by Fischhoff & MacGregor (19) provides another example of the dangers of relying on a single response mode to describe behavior. They used four different response modes to ask about the chances of dying, given that one was afflicted with each of various maladies (e.g. how many people die out of each 100,000 who get influenza; how many people died out of the 80 million who caught influenza last year). Again, there was strong internal consistency across response modes, whereas absolute estimates varied over as much as two orders of magnitude. A follow-up study reduced this range by providing an independent basis for eliminating the response mode that produced the most discrepant results (e.g. subjects were least able to remember statistics reported in that format—estimating the number of survivors for each person who succumbed to a problem).

PERCEIVED INVULNERABILITY Estimating the accuracy of risk estimates requires not only an appropriate response mode, but also credible statistical estimates against which responses can be compared. The studies just described asked about population risks in situations where credible statistical estimates were available. Performance might be different (poorer?) for risks whose magnitude is less readily calculated. Furthermore, people may not see these population risks as personally relevant.

As a partial way to avoid these problems, some investigators have asked subjects to judge whether they are more or less at risk than others in more or less similar circumstances (63, 69). They find that most people in most situations see themselves as facing less risk than average others (which could, of course, be true for only half a population). A variety of processes could account for such a bias, including both cognitive ones (e.g. the greater availability of the precautions that one takes) and motivational ones (e.g. wishful thinking). To the extent that this bias exists in the world outside the

experiment and interview, such a bias could prompt unwanted risk taking (e.g. because warnings seem more applicable to other people).[4]

Defining Risk

These studies attempt to measure risk perceptions under the assumption that people define "risk" as the probability of death. Anecdotal observation of scientific practice shows that "risk" means different things in different contexts (8, 23). For some analysts, risk is expected loss of life of expectancy; for others, it is expected probability of premature fatality (with the former definition placing a premium on deaths among the young). Some of the apparent disagreement between experts and laypeople regarding the magnitude of risks in society may be due to differing definitions of risk (20, 62).

CATASTROPHIC POTENTIAL One early study asked experts and laypeople to estimate the "risk of death" faced by society as a whole from 30 activities and technology (62). The experts' judgments could be predicted well from statistical estimates of average-year fatalities—as could the estimates of laypeople given that specific definition. Lay estimates of "risk" were more poorly correlated with average-year fatalities. However, much of the residual variance could be predicted by their estimates of catastrophic potential, the ability to cause large numbers of death in a nonaverage year. Thus, casual observation had obscured the extent to which experts and laypeople agreed about routine death tolls (for which scientific estimates are relatively uncontroversial) and disagreed about the possibility of anomalies (for which the science is typically weaker).

Sensing that there was something special about catastrophic potential, some risk experts have suggested that social policy give extra weight to hazards carrying that kind of threat. One experimental study has, however, found that people may not care more for many lives lost in a single accident than for the same number of lives lost in separate incidents (61).[5] The critical factor in catastrophic potential is not how the deaths are grouped, but the possibility of discovering that a technology is out of control. Such "surprise potential"

[4]In a recent study (53), we derived judgments of relative risk from judgments of the absolute degree of risk that people assigned to themselves and to target others (a close friend, an acquaintance, a parent, a child). On a response scale that facilitated expressing very low probabilities, subjects assigned a probability of less than 1 in 10 million about 10% of the time and a probability of less than 1 in 10,000 about one third of the time. The events involved "a death or injury requiring hospitalization over the next five years" from sources like auto accidents, drug addiction, and explosions. Here, too, middle-class adults and adolescents responded similarly, despite the common belief that teens take risks, in part, because of a unique perception of invulnerability (11).

[5]When accidents involving large numbers of fatalities are easy to imagine, catastrophic potential can be rated high because of availability, even when estimates of average-year fatalities are relatively low, as was the case for nuclear power in this study.

is strongly correlated with catastrophic potential in people's judgments (and, presumably, in scientific estimates). However, the two features represent rather different ethical bases for distinguishing among risks.

DIMENSIONS OF RISK Recognizing that correlated features can confuse the interpretation of risk behaviors, investigators have looked extensively at the patterns of correlations among features (1, 22, 60). Overall, they have found a remarkably robust picture, typically revealing two or three dimensions of risk, which capture much of the variation in judgments of up to 20 aspects of risk. The general structure of this "risk space" is relatively similar across elicitation method, subject population (e.g. experts versus laypeople), and risk domain. Core concepts in these dimensions include how well a risk is understood and how much of a feeling of dread it evokes. The placement of individual hazards in the space does vary with individual and with group, in ways that can predict judgments of risk management policies (e.g. how tightly a technology should be regulated). Relatively little is known about the role of these dimensions in individual risk decisions.

RISK COMPARISONS The multidimensional character of risk means that hazards that are similar in many ways may still evoke quite different responses. This fact is neglected in appeals to accept one risk, because one has accepted another that is similar to it in some ways (8, 18). The most ambitious of these appeals present elaborate lists of hazards, the exposure to which is adjusted so that they pose equivalent risks (e.g. both one tablespoon of peanut butter and 50 years of living at the boundary of a nuclear power plant create a one-in-a-million risk of premature death). Recognizing that such comparisons are often perceived as self-serving, the Chemical Manufacturers Association (6) commissioned a guide to risk comparisons, which presents such lists, but with the attached caution, WARNING! USE OF DATA IN THIS TABLE FOR RISK COMPARISON PURPOSES CAN DAMAGE YOUR CREDIBILITY.[6]

QUALITATIVE ASSESSMENT

Event Definitions

Scientific estimates of risk require detailed specification of the conditions under which it is to be observed. For example, a fertility counselor estimating a woman's risk of an unplanned pregnancy would consider the amount of

[6]The guide also offers advice on how to make risk comparisons, if one feels the compulsion, along with examples of more and less acceptable comparisons. Although the advice is logically derived from risk perception research, it was not tested empirically. In such a test, we found little correlation between the predicted degree of acceptability and the acceptability judgments of several diverse groups of subjects (56).

intercourse, the kinds of contraceptive used (and the diligence with which they are applied), her physiological condition (and that of her partner), and so on. If laypeople are to make accurate assessments, then they require the same level of detail. That is true whether they are estimating risks for their own sake or for the benefit of an investigator studying risk perceptions.

When such investigators omit needed details, they create adverse conditions for subjects. To respond correctly, subjects must first guess the question and then know the answer to it. Consider, for example, the question, "What is the probability of pregnancy with unprotected sex?" A well-informed subject who understood this to mean a single exposure would be seen as underestimating the risk by an investigator who intended the question to mean multiple exposures.

Such ambiguous events are common in surveys designed to study public perceptions of risk. For example, a National Center for Health Statistics survey (70) question asked, "How likely do you think it is that a person will get the AIDS virus from sharing plates, forks, or glasses with someone who had AIDS?" Even if the survey had not used an ambiguous response mode (very likely, unlikely, etc.), it would reveal relatively little about subjects' understanding of disease risks. For their responses to be meaningful, subjects must spontaneously assign the same value to each missing detail, while investigators guess what subjects decided.

We asked a relatively homogeneous group of subjects what they thought was meant regarding the amount and kind of sharing implied by this question (after they had answered it) (16). These subjects generally agreed about the kind of sharing (82% interpreted it as sharing during a meal), but not about the amount (a single occasion, 39%; several occasions 20%; routinely, 28%; uncertain, 12%). A survey question about the risks of sexual transmission evoked similar disagreement. We did not study what readers of the survey's results believed about subjects' interpretations.

Supplying Details

Aside from their methodological importance, the details that subjects infer can be substantively interesting. People's intuitive theories of risk are revealed in the variables that they note and the values that they supply. In a systematic evaluation of these theories, Quadrel (52) asked adolescents to think aloud as they estimated the probability of several deliberately ambiguous events (e.g. getting in an accident after drinking and driving, getting AIDS through sex).

These subjects typically wondered (or made assumptions) about numerous features. In this sense, subjects arguably showed more sophistication than the investigators who created the surveys from which these questions were taken or adapted. Generally speaking, these subjects wre interested in variables that could figure in scientific risk analyses (although scientists might not yet know

what role each variable plays). There were, however, some interesting exceptions. Although subjects wanted to know the "dose" involved with most risks, they seldom asked about the amount of sex in one question about the risks of pregnancy and in another question about the risks of HIV transmission. They seemed to believe that an individual either is or is not sensitive to the risk, regardless of the amount of the exposure. In other cases, subjects asked about variables without a clear connection to risk level (e.g. how well members of the couple knew one another).

In a follow-up study, Quadrel (52) presented richly specified event descriptions to teens drawn from the same populations (school organizations and substance abuse treatment homes). Subjects initially estimated the probability of a risky outcome on the basis of some 20 details. Then, they were asked how knowing each of three additional details would change their estimates. One of those details had been provided by subjects in the preceding study; two had not. Subjects in this study responded to the relevant detail much more than to the irrelevant ones. Thus, at least in these studies, teens did not balk at making judgments regarding complex stimuli and revealed consistent intuitive theories in rather different tasks.

Cumulative Risk—A Case in Point

As knowledge accumulates about people's intuitive theories of risk, it will become easier to predict which details subjects know and ignore, as well as which omissions they will notice and rectify. In time, it might become possible to infer the answers to questions that are not asked from answers to ones that are—as well as the inferences that people make from risks that are described explicitly to risks that are not. The invulnerability results reported above show the need to discipline such extrapolations with empirical research. Asking people about the risks to others like themselves is not the same as asking about their personal risk. Nor need reports about others' risk levels be taken personally.

One common, and seemingly natural, extrapolation is between varying numbers of independent exposures to a risk. Telling people the risk from a single exposure should allow them to infer the risk from whatever multiple they face; asking subjects what risk they expect from one amount should allow one to infer what they expect from other amounts. Unfortunately, for both research and communication, teens' insensitivity to the amount of intercourse (in determining the risks of pregnancy or HIV transmission) proves to be a special case of a general problem. Several reviews (9, 48) have concluded that between one third and one half of sexually active adolescents explain not using contraceptives with variants of, "I thought I (or my partner) couldn't get pregnant." Another study (59) found that adults greatly underestimated

the rate at which the risk of contraceptive failure accumulates through repeated exposure, even after eliminating (from the data analysis) the 40% or so of subjects who saw no relationship between risk and exposure. One corollary of this bias is not realizing the extent to which seemingly small differences in annual failure rates (what is typically reported) can lead to large differences in the cumulative risk associated with continued use.

After providing practice with a response mode designed to facilitate the expression of small probabilities, Linville et al (41) asked college students to estimate the risks of HIV transmission from a man to a woman as the result of 1, 10, or 100 cases of protected sex. For one contact, the median estimate was .10, a remarkably high value according to public health estimates (14, 33). For 100 contacts, however, the median estimate was .25, a more reasonable value. Very different pictures of people's risk perceptions would emerge from studies that asked just one of these questions or the other. Risk communicators could achieve quite different effects if they chose to describe the risk of one exposure and not the other. They might create confusion if they chose to communicate both risks, thus leaving recipients to reconcile the seeming inconsistency.

Mental Models of Risk Processes

THE ROLE OF MENTAL MODELS These intuitive theories of how risks accumulate were a byproduct of research intended to improve the elicitation and communication of quantitative probabilities. Such research can serve the interests of individuals who face well-formulated decisions in which estimates of health risks (or benefits) play clearly defined roles. For example, a homeowner poised to decide whether to test for radon needs estimates of the cost and accuracy of tests, the health risks of different radon levels, the cost and efficacy of ways to mitigate radon problems, and so on (64).

Often, however, people are not poised to decide anything. Rather, they just want to know what the risk is and how it works. Such substantive knowledge is essential for following an issue in the news media, for participating in public discussions, for feeling competent to make decisions, and for generating options among which to decide. In these situations, people's objective is to have intuitive therories that correspond to the main elements of the reigning scientific theories (emphasizing those features relevant to control strategies).

The term mental model is often applied to intuitive theories that are elaborated well enough to generate predictions in diverse circumstances (24). Mental models have a long history in psychology (7, 50). For example, they have been used to examine how people understand physical processes (26),

international tensions (43), complex equipment (57), energy conservation (34), and the effects of drugs (31).

If these mental models contain critical bugs, they can lead to erroneous conclusions, even among otherwise well-informed people. For example, not knowing that repeated sex increases the associated risks could undermine much other knowledge. Bostrom et al (5) found that many people know that radon is a colorless, odorless, radioactive gas. Unfortunately, some also associate radioactivity with permanent contamination. However, this widely publicized property of high-level waste is not shared by radon. Not realizing that the relevant radon byproducts have short half-lives, homeowners might not even bother to test (believing that there was nothing that they could do, should a problem be detected).

ELICITING MENTAL MODELS In principle, the best way to detect such misconceptions would be to capture people's entire mental model on a topic. Doing so would also identify those correct conceptions upon which communications could build (and which should be reinforced). The critical threat to capturing mental models is reactivity, i.e. changing respondents as a result of the elicitation procedure. One wants neither to induce nor to dispell misconceptions, either through leading questions or subtle hints. The interview should neither preclude the expression of unanticipated beliefs nor inadvertently steer subjects around topics (13, 24, 28).

Bostrom et al (5) offer one possible compromise strategy, which has been used for a variety of risks (2, 42, 47). Their interview protocol begins with very open-ended questions: They ask subjects what they know about a topic, then prompt them to consider exposure, effects, and mitigation issues. Subjects are asked to elaborate on every topic mentioned. Once these minimally structured tasks are exhausted, subjects sort a large stack of diverse photographs, according to whether each seems related to the topic, and explain their reasoning as they go.

Once transcribed, the interviews are coded into an expert model of the risk. This is a directed network, or influence diagram (29), which shows the different factors affecting the magnitude of the risk. The expert model is created by iteratively pooling the knowledge of a diverse group of experts. It might be thought of as an expert's mental model, although it would be impressive for any single expert to produce it all in a single session following the open-ended interview protocol. Figure 2 shows the results of coding one subject's interview into the expert model for radon. The subject's concepts were characterized as correct, incorrect, peripheral (technically correct, but only distantly related to the topic), background (referring to general principles of science), evaluative, and nonspecific (or vague).

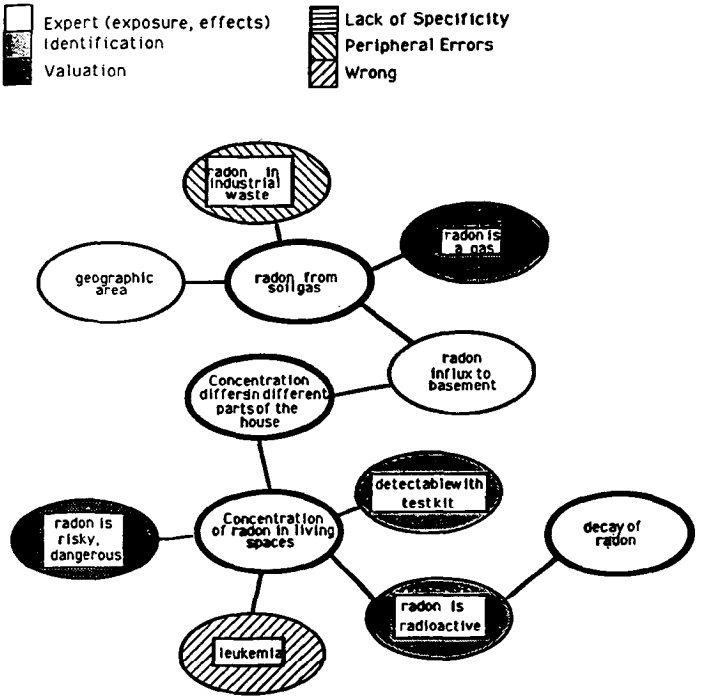

Figure 2 One subject's model of processes affecting radon risk, elicited with an open-ended interview protocol. (From Ref. 5.)

CREATING COMMUNICATIONS

Selecting Information

The first step in designing communications is to select the information that they should contain. In many existing communications, this choice seems arbitrary, reflecting some expert or communicator's notion of "what people ought to know." Poorly chosen information can have several negative consequences, including both wasting recipients' time and being seen to waste it (thereby reflecting insensitivity to their situation). In addition, recipients will be judged unduly harshly if they are uninterested in information that, to them, seems irrelevant. The Institute of Medicine's fine and important report, *Confronting AIDS* (30), despaired after a survey showed that only 41% of the public knew that AIDS was caused by a virus. Yet, one might ask what role that information could play in any practical decision (as well as what those subjects who answered correctly meant by "a virus").

The information in a communication should reflect a systematic theoretical perspective, capable of being applied objectively. Here are three candidates for such a perspective, suggested by the research cited above:

MENTAL MODEL ANALYSIS Communications could attempt to convey a comprehensive picture of the processes creating (and controlling) a risk. Bridging the gap between lay mental models and expert models would require adding missing concepts, correcting mistakes, strengthening correct beliefs, and deemphasizing peripheral ones.

CALIBRATION ANALYSIS Communications could attempt to correct the critical "bugs" in recipients' beliefs. These are defined as cases where people confidently hold incorrect beliefs that could lead to inappropriate actions (or lack enough confidence in correct beliefs to act on them).

VALUE-OF-INFORMATION ANALYSIS Communications could attempt to provide the pieces of information that have the largest possible impact on pending decisions. Value-of-information analysis is the general term for techniques that determine the sensitivity of decisions to different information (46).

The choice among these approaches would depend on, among other things, how much time is available for communication, how well the decisions are formulated, and what scientific risk information exists. For example, calibration analysis might be particularly useful for identifying the focal facts for public service announcements. Such facts might both grab recipients' attention and change their behavior. A mental model analysis might be more suited for the preparation of explanatory brochures or curricula.

Merz (45) applied value-of-information analysis to a well-specified medical decision, whether to undergo carotid endarterectomy. Both this procedure, which involves scraping out an artery that leads to the head, and its alternatives have a variety of possible positive and negative effects. These effects have been the topic of extensive research, which has provided quantitative risk estimates of varying precision. Merz created a simulated population of patients, who varied in their physical condition and relative preferences for different health states. He found that knowing about a few, but only a few, of the possible side effects would change the preferred decision for a significant portion of patients. He argued that communications focused on these few side effects would make better use of patients' attention than laundry lists of undifferentiated possibilities. He also argued that his procedure could provide an objective criterion for identifying the information that must be transmitted to insure medical informed consent.

Formatting Information

Once information has been selected, it must be presented in a comprehensible way. That means taking into account the terms that recipients use for understanding individual concepts and the mental models they use for integrating those concepts. It also means respecting the results of research into text comprehension. That research shows, for example, that comprehension improves when text has a clear structure and, especially, when that structure conforms to recipients' intuitive representation of a topic; that critical information is more likely to be remembered when it appears at the highest level of a clear hierarchy; and that readers benefit from "adjunct aids," such as highlighting, advanced organizers (showing what to expect), and summaries. Such aids might even be better than full text for understanding, retaining, and being able to look up information. Fuller treatment than is possible here can be found in Refs. 12, 25, 35, 54, 58.

There may be several different formats that meet these general constraints. Recently, we created two brochures that presented clear but different structures to explain the risks of radon (4). One was organized around a decision tree, which showed the options facing home owners, the probabilities of possible consequences, and the associated costs and benefits. The second was organized around a directed network, in effect, the expert model of the mental model studies. Both were compared with the Environmental Protection Agency's (EPA) widely distributed (and, to EPA's great credit, heavily evaluated) *Citizen's Guide to Radon* (65a), which uses primarily a question-and-answer format, with little attempt to summarize or impose a general structure. All three brochures substantially increased readers' understanding of the material presented in them. However, the structured brochures did better (and similar) jobs of enabling readers to make inferences about issues not mentioned explicitly and to give explicit advice to others.

Evaluating Communications

Effective risk communications can help people to reduce their health risks, or to get greater benefits in return for those risks that they take. Ineffective communications not only fail to do so, but also incur opportunity costs, in the sense of occupying the place (in recipients' lives and society's functions) that could be taken up by more effective communications. Even worse, misdirected communications can prompt wrong decisions by omitting key information or failing to contradict misconceptions, create confusion by prompting inappropriate assumptions or by emphasizing irrelevant information, and provoke conflict by eroding recipients' faith in the communicator. By causing undue alarm or complacency, poor communications can have

greater public health impact than the risks that they attempt to describe. It may be no more acceptable to release an untested communication than an untested drug. Because communicators' intuitions about recipients' risk perceptions cannot be trusted, there is no substitute for empirical validation (17, 20, 49, 55, 60).

The most ambitious evaluations ask whether recipients follow the recommendations given in the communication (37, 68). However, that standard requires recipients not only to understand the message, but also to accept it as relevant to their personal circumstances. For example, home owners without the resources to address radon problems might both understand and ignore a communication about testing; women might hear quite clearly what an "expert" is recommending about how to reduce their risk of sexual assault, yet reject the political agenda underlying that advice (15). Judging the effectiveness of a program by behavioral effects requires great confidence that one knows what is right for others.

A more modest, but ethically simpler, evaluation criterion is ensuring that recipients have understood what a message was trying to say. That necessary condition might prove sufficient, too, if the recommended action is, indeed, obviously appropriate, once the facts are known. Formal evaluations of this type seem to be remarkably rare, among the myriad of warning labels, health claims and advisories, public service announcements, and operating instructions that one encounters in everyday life and work.

Evaluating what people take away from communications faces the same methodological challenges as measuring ambient risk perceptions. To elaborate slightly on a previous section, the evaluator wants to avoid reactivity, changing people's beliefs through the cues offered by how questions and answers are posed; illusory expertise, restricting the expression of inexpert beliefs; and illusory discrimination, suppressing the expression of inconsistent beliefs.

For example, as part of an ambitious program to evaluate its communications regarding the risks of radon, the EPA (10) posed the following question: "What kinds of problems are high levels of radon exposure likely to cause? a. minor skin problems; b. eye irritations; c. lung cancer." This question seems to risk inflating subjects' apparent level of understanding in several ways. Subjects who know only that radon causes cancer might deduce that it causes lung cancer. The words "minor" and "irritation" might imply that these are not the effects of "high levels" (of anything). There is no way to express other misconceptions, such as that radon causes breast cancer and other lung problems, which emerged with some frequency in our open-ended interviews (5).

In principle, open-ended interviews provide the best way to reduce such

threats. However, they are very labor intensive. The stakes riding on many risk communications might justify that investment. Realistically speaking, the needed time and financial resources are not always available. As a result, open-ended, one-on-one interviews are better seen as necessary stepping stones to structured questionnaires, suitable for mass administration. Those questionnaires should cover the critical topics in the expert model, express questions in terms familiar to subjects, and test for the prevalence of misconceptions. Worked examples can be found in Ref. 4.

CONCLUSION

Risk perception and risk communication research are complicated businesses, perhaps as complicated as assessing the magnitude of the risks that they consider. A chapter of this length can, at best, indicate the dimensions of complexity and the directions of plausible solutions. In this treatment, we have emphasized methodological issues because we believe that these topics often seem deceptively simple to those not trained in them. Because we all talk and ask questions in everyday life, it seems straightforward to do so regarding health risks. Unfortunately, there are many pitfalls to such amateurism, hints to which can be found in those occasions in life where we have misunderstood or been misunderstood, particularly when dealing with strangers on unfamiliar topics.

Research in this area is fortunate in being able to draw on well-developed literatures in such areas as cognitive, health, and social psychology; survey research; psycholinguistics; psychophysics; and behavioral decision theory. It is unfortunate in having to face the particularly rigorous demands of assessing and improving beliefs about health risks. These often involve complex and unfamiliar topics, surrounded by unusual kinds of uncertainty, for which individuals and groups lack stable vocabularies. Health risk decisions also raise difficult and potentially threatening tradeoffs. Even the most carefully prepared and evaluated communications may not be able to eliminate the anxiety and frustration that such decisions create. However, systematic preparation can keep communications from adding to the problem. At some point in complex decisions, we "throw up our hands" and go with what seems right. Good risk communications can help people get further into the problem before that happens.

Health risk decisions are not just about cognitive processes and cooly weighed information. Emotions play a role, as do social processes. Nonetheless, it is important to get the cognitive part right, lest people's ability to think their way to decisions be underestimated and underserved.

Literature Cited

1. Arabie, P., Maschmeyer, C. 1988. Some current models for the perception and judgment of risk. *Org. Behav. Hum. Decis. Proc.* 41:300–29
2. Atman, C. 1990. *Network structures as a foundation for risk communication*. Doctoral diss. Carnegie Mellon Univ.
3. Beyth-Marom, R. 1982. How probable is probable? Numerical translation of verbal probability expressions. *J. Forecasting* 1:257–69
4. Bostrom, A., Atman, C., Fischhoff, B., Morgan, M. B. 1993. Evaluating risk communications: completing and correcting mental models of hazardous processes. Submitted.
5. Bostrom, A., Fischhoff, B., Morgan, M. G. Eliciting mental models of hazardous processes: a methodology and an application to radon. *J. Soc. Issues*. In press
6. Covello, V. T., Sandman, P. M., Slovic, P. 1988. *Risk Communication, Risk Statistics, and Risk Comparisons: A Manual for Plant Managers*. Washington, DC: Chem. Manuf. Assoc.
7. Craik, K. 1943. *The Nature of Explanation*. Cambridge: Cambridge Univ. Press
8. Crouch, E. A. C., Wilson, R. 1982. *Risk/Benefit Analysis*. Cambridge, Mass: Ballinger
9. Cvetkovich, G., Grote, B., Bjorseth, A., Sarkissian, J. 1975. On the psychology of adolescents' use of contraceptives. *J. Sex Res.* 11:256–70
10. Desvousges, W. H., Smith, V. K., Rink, H. H. III. 1989. *Communicating Radon Risk Effectively: Radon Testing in Maryland*. EPA 230-03-89-408. Washington, DC: US Environ. Prot. Agency, Off. Policy, Plan. Eval.
11. Elkind, D. 1967. Egocentrism in adolescence. *Child Dev.* 38:1025–34
12. Ericsson, K. A. 1988. Concurrent verbal reports on text comprehension: a review. *Text* 8:295–325
13. Ericsson, K. A., Simon, H. A. 1980. Verbal reports as data. *Psychol. Rev.* 87:215–51
14. Fineberg, H. V. 1988. Education to prevent AIDS. *Science* 239:592–96
15. Fischhoff, B. 1992. Giving advice: decision theory perspectives on sexual assault. *Am. Psychol.* 47:577–88
16. Fischhoff, B. 1989. Making decisions about AIDS. In *Primary Prevention of AIDS*, ed. V. Mays, G. Albee, S. Schneider, pp. 168–205. Newbury Park, Calif: Sage
17. Fischhoff, B. 1987. Treating the public with risk communications: a public health perspective. *Sci. Technol. Hum. Values* 12:13–19
18. Fishhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L., Keeney, R.L. 1981. *Acceptable Risk*. New York: Cambridge Univ. Press
19. Fischhoff, B., MacGregor, D. 1983. Judged lethality: how much people seem to know depends upon how they are asked. *Risk Anal.* 3:229–36
20. Fischhoff, B., Slovic, P., Lichtenstein, S., 1983. The "public" vs. the "experts": perceived vs. actual disagreement about the risks of nuclear power. In *Analysis of Actual vs. Perceived Risks*, ed. V. Covello, G. Flamm, J. Rodericks, R. Tardiff, pp. 235–49. New York: Plenum
21. Fischhoff, B., Slovic, P., Lichtenstein, S., 1977. Knowing with certainty: the appropriateness of extreme confidence. *J. Exp. Psychol. Hum. Percept. Perform.* 3:552–64
22. Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., Combs, B. 1978. How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sci.* 8:127–52
23. Fischhoff, B., Watson, S., Hope, C. 1984. Defining risk. *Policy Sci.* 17: 123–39
24. Galotti, K. M. 1989. Approaches to studying formal and everyday reasoning. *Psychol. Bull.* 105:331–51
25. Garnham, A. 1987. *Mental Models as Representations of Discourse and Text*. New York: Halsted
26. Gentner, D., Stevens, A. L., eds. 1983. *Mental Models*. Hillsdale, NJ: Erlbaum
27. Heimer, C. A. 1988. Social structure psychology, and the estimation of risk. *Annu. Rev. Soc.* 14:491–519
28. Hendrickx, L. C. W. P. 1991. *How versus how often: the role of scenario information and frequency information in risk judgment and risky decision making*. Doctoral Diss., Rijksuniversiteit Groningen
29. Howard, R. A. 1989. Knowledge maps. *Manage. Sci.* 35:903–22
30. Inst. of Med. 1986. *Confronting AIDS*. Washington, DC: Natl. Acad. Press
31. Jungermann, H., Shutz, H., Thuring, M. 1988. Mental models in risk assessment: informing people about drugs. *Risk Anal.* 8:147–55

32. Kahneman, D., Slovic, P., Tversky, A., eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge Univ. Press
33. Kaplan, E. H. 1989. What are the risks of risky sex? *Oper. Res.* 37:198–209
34. Kempton, W., 1987. Variation in folk models and consequent behavior. *Am. Behav. Sci.* 31:203–18
35. Kintsch, W. 1986. Learning from text. *Cogn. Instr.*3:87–108
36. Krimsky, S., Plough, A. 1988. *Environmental Hazards*. Dover, Mass: Auburn
37. Lau, R., Kaine, R., Berry, S., Ware, J., Roy, D. 1980. Channeling health: a review of the evaluation of televised health campaigns. *Health Educ. Q.* 7:56–89
38. Lichtenstein, S., Fischhoff, B., Phillips, L. D. 1982. Calibration of probabilities: state of the art to 1980. See Ref. 32
39. Lichtenstein, S., Newman, J. R. 1967. Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychon. Sci.* 9:563–64
40. Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., Combs, B. 1978. Judged frequency of lethal events. *J. Exp. Psychol. Hum. Learn. Mem.* 4: 551–78
41. Linville, P. W., Fischer, G. W., Fischhoff, B. Perceived risk and decision making involving AIDS. 1993. In *The Social Psychology of HIV Infection*, ed. J. B. Pryor, G. D. Reeder. Hillsdale, NJ: Erlbaum. In press
42. Maharik, M., Fischhoff, B. 1992. The risks of nuclear energy sources in space: Some activists' perceptions. *Risk Anal.* 12:383–92
43. Means, M. L., Voss, J. F., 1985. Star wars: a developmental study of expert and novice knowledge studies. *J. Mem. Lang.* 24:746–57
44. Merton, R. F. 1987. The focussed interview and focus groups. *Public Opin. Q.* 51:550–66
45. Merz, J. F. 1991. *Toward a standard of disclosure for medical informed consent: development and demonstration of a decision—analytic methodology*. Ph.D. diss. Carnegie Mellon Univ.
46. Merz, J. F., Fischhoff, B., 1990. Informed consent does not mean rational consent: cognitive limitations on decision-making. *J. Legal Med.* 11: 321–50
47. Morgan, M. G., Florig, H. K., Nair, I., Cortes, C., Marsh, K., Pavlosky, K. 1990. Lay understanding of power-frequency fields. *Bioelectromagnetics* 11:313–35
48. Morrison, D. M. 1985. Adolescent contraceptive behavior: a review. *Psychol. Bull.* 98:538–68
49. Natl. Res. Counc. 1989. *Improving Risk Communication*. Washington, DC: Natl. Acad. Press
50. Oden, G. C. 1987. Concept, knowledge, and thought. *Annu. Rev. Psychol.* 38:203–27
51. Poulton, E. C. 1989. *Bias in Quantifying Judgment*. Hillsdale, NJ: Erlbaum
52. Quadrel, M. J. 1990. *Elicitation of adolescents' perceptions: qualitative and quantitative dimensions*. Ph.D. diss. Carnegie Mellon Univ.
53. Quadrel, M. J., Fischhoff, B., Davis, W. 1993. Adolescent invulnerability. *Am. Psychol.* In press
54. Reder, L. M. 1985. Techniques available to author, teacher, and reader to improve retention of main ideas of a chapter. In *Thinking and Learning Skills: Vol. 2. Research and Open Questions*, ed. S. F. Chipman, J. W. Segal, R. Glaser, pp. 37–64. Hillsdale, NJ: Erlbaum
55. Rohrmann, B., 1990. *Analyzing and evaluating the effectiveness of risk communication programs*. Unpubl. Univ. of Mannheim
56. Roth, E., Morgan, G., Fischhoff, B., Lave, L., Bostrom, A. 1990. What do we know about making risk comparisons? *Risk Anal.* 10:375–87
57. Rouse, W. B., Morris, N. M. 1986. On looking into the black box: prospects and limits in the search for mental models. *Psychol. Bull.* 100:349–63
58. Schriver, K. A. 1989. *Plain language for expert or lay audiences: designing text using protocol aided revision*. Pittsburgh: Commun. Design Cent., Carnegie Mellon Univ.
59. Shaklee, H., Fischhoff, B. 1990. The psychology of contraceptive surprises: judging the cumulative risk of contraceptive failure. *J. Appl. Psychol.* 20: 385–403
60. Slovic, P. 1987. Perceptions of risk. *Science* 236:280–85
61. Slovic, P., Fischhoff, B., Lichtenstein, S. 1984. Modeling the societal impact of fatal accidents. *Manage Sci.* 30:464–74
62. Slovic, P., Fischhoff, B., Lichtenstein, S. 1979. Rating the risks. *Environment* 21:14–20, 36–39
63. Svenson, O. 1981. Are we all less risky and more skillful than our fellow drivers? *Acta Psychol.* 47:143–48
64. Svenson, O., Fischhoff, B. 1985. Lev-

els of environmental decisions. *J. Environ. Psychol.* 5:55–68
65. Tversky, A., Kahneman, D. 1973. Availability: a heuristic for judging frequency and probability. *Cogn. Psychol.* 4:207–32
65a. US Environ. Prot. Agency Off. of Air and Radiat., US Dep. of Health and Hum. Serv., Cent. for Dis. Control. 1986. *A Citizen's Guide to Radon: What It Is and What to Do About It,* OPA-86-004
66. von Winterfeldt, D., Edwards, W., 1986. *Decision Analysis and Behavioral Research*. New York: Cambridge Univ. Press
67. Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R., Forsyth, B. 1986. Measuring the vague meanings of probability terms. *J. Exp. Psychol. Gen.* 115:348–65
68. Weinstein, N. 1987. *Taking Care: Understanding and Encouraging Self-Protective Behavior*. New York: Cambridge Univ. Press
69. Weinstein, N. D. 1980. Unrealistic optimism about future life events. *J. Pers. Soc. Psychol.* 39:806–20
70. Wilson, R. W., Thornberry, O. T. 1987. Knowledge and attitudes about AIDS: provisional data from the National Health Interview Survey, August 10–30, 1987. *Adv. Data,* No. 146
71. Yates, J. F., ed. 1992. *Risk Taking*. Chichester: Wiley
72. Yates, J. F. 1989. *Judgment and Decision Making*. Englewood Cliffs, NJ: Prentice Hall

Annual Review of Public Health
Volume 14, 1993

CONTENTS