Oxford Medicine

Oxford Textbook of Global Public Health (6 ed.)

Edited by Roger Detels, Martin Gulliford, Quarraisha Abdool Karim, and Chorh Chuan Tan

Publisher: Oxford University Press

Print ISBN-13: 9780199661756

DOI: 10.1093/med/9780199661756.001.0001

Print Publication Date: Feb 2015

Published online: Feb 2015

Risk perception and communication

Chapter: Risk perception and communication

Author(s): Baruch Fischhoff

DOI: 10.1093/med/9780199661756.003.0138

Introduction to risk perception and communication

Sound health risk decisions require understanding the risks and benefits of possible actions. Some of those choices are personal. They include whether to wear bicycle helmets and seat belts, whether to read and follow safety warnings, whether to buy and use condoms, and how to select and cook food. Other choices are made as citizens. They include whether to protest the siting of hazardous waste incinerators and halfway houses, whether to support fluoridation and 'green' candidates, and what to include in sex education.

Sometimes, single choices have large effects (e.g. buying a safe car, taking a dangerous job, getting pregnant). Sometimes, small effects accumulate over multiple choices (e.g. exercising, avoiding trans-fats, wearing seatbelts, using escort services). Sometimes, health-related choices focus on health; sometimes, they do not (e.g. purchasing homes that require long commutes, choosing friends who exercise regularly, joining religious groups opposed to vaccination).

This chapter reviews the research base for assessing and improving individuals' understanding of the risks and possible benefits of health-related choices. Following convention, these pursuits are called *risk perception* and *risk communication*, respectively, even though the same basic behavioural principles apply to the benefits that all risk decisions entail, if only the benefits of reducing risks (Fischhoff et al. 2011). Psychologists sometimes reserve the term 'perception' for direct physiological responses to stimuli, using 'judgement' for the translation of those responses into observable estimates. A perennial research topic is identifying the conditions under which judgement surrenders to perception, and when emotions play little role because people know what they want to do (Slovic et al. 2005). This chapter emphasizes judgement, hoping to expand the envelope of deliberative processes in personal and public health decisions.

Inaccurate judgements about risks can harm people. So can inaccurate beliefs about those judgements. If their understanding is overestimated, then people may face impossibly hard choices (e.g. among unfamiliar medical alternatives, without adequate counselling). If their understanding is underestimated, then people may be needlessly denied the right to choose. As a result, the chapter assumes: (1) that descriptive statements about people's beliefs must be underpinned by empirical evidence and (2) that evaluative statements about the adequacy of people's understanding must be founded on rigorous analysis of what they need to know, in order to make a sound decision. To these ends, the chapter emphasizes methodological safeguards against misguided assessments.

The next section, 'Quantitative assessment', considers beliefs about how large risks are. The following section, 'Qualitative assessment', treats beliefs about the processes that create and control risks, on the basis of which people produce and evaluate quantitative estimates. Both sections address both measurement issues and barriers to understanding. The next section, 'Creating communications,' provides a structured approach for developing communications about health-related decisions, focused on individuals' information needs. The 'Conclusion' section considers the strategic importance of risk communication in public health. Access to research on complementary social and emotional processes might begin with Breakwell (2007), Krimsky and Golding (1992), and Peters and McCaul (2005).

Quantitative assessment

Estimating risk magnitude

A common complaint among experts is that 'the public doesn't realize how small (or large) Risk X is'. There is empirical evidence demonstrating such biases (Slovic 2001). However, that evidence has often been collected in settings designed to reveal biases. Looking for problems is a standard strategy in experimental sciences, designed to reveal the processes creating those problems, but not their prevalence or magnitude in specific domains of everyday life. Generalizing from research decisions to real-world ones requires matching the conditions in each. Looking at the details of one widely cited study shows how that matching process might proceed, while introducing some general principles and results.

Participants

Lichtenstein et al. (1978) asked members of a civic group in Eugene, Oregon, to estimate the

annual number of deaths in the United States from 30 causes (e.g. botulism, tornadoes, motor vehicle accidents). They were older than the college students often studied by psychologists. Age could affect *what* people think, as a result of differences in their education and life experience. It is less likely to affect *how* they think. Many cognitive processes are widely shared, once people pass middle adolescence, unless they suffer some impairment (Fischhoff 2008; Reyna and Farley 2006; Finucane and Gullion 2010).

One widely shared class of cognitive processes relies on judgemental *heuristics* to infer unknown quantities (Kahneman et al. 1982; Gilovich et al. 2003). One well-known heuristic is *availability*, whereby people assess an event's probability by how easily instances come to mind. Although more available events are often more likely, media coverage (among other things) makes some events disproportionately available, thereby inducing biased judgements —unless people take into account how appearances can be deceiving. *How* people generate instances of events, using their memory and imagination, should reflect widely shared general cognitive processes. *What* those memories and images contain, as well as what faith people place in information sources, should vary with their experiences.

Lichtenstein et al. (1978) elicited judgements with two *response modes*. One asked people to pick the more frequent of two paired causes of death (e.g. asthma, botulism) and then to estimate the ratio of their frequencies. The second asked for the number of deaths, after providing the value for one cause (either electrocution or motor vehicle accidents) in order to give respondents a feeling for annual death rates—after pretests found that many people had little idea about what range of numbers to give. Fig. **7.6.1** shows results with the second method, which are typical of such studies.
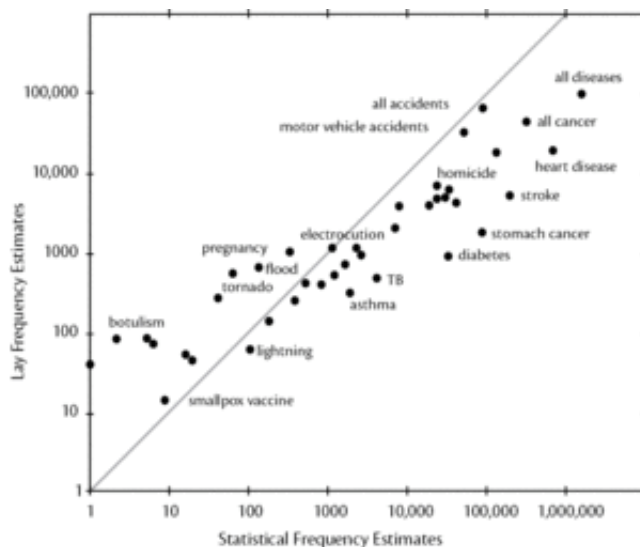

Fig. 7.6.1
Best quadratic fit line to geometric mean judgements of the annual toll from 40 causes of death in the United States, compared to best available statistical estimates.

Reproduced from Fischhoff, B. and Kadvany, J., *Risk: A Very Short Introduction,* Figure 12, p. 92, Oxford University Press, Oxford, UK, Copyright © Baruch Fischhoff and John Kadvany 2011, by permission of Oxford University Press.

Results

1. Judgements of the *relative* risk from different causes were similar however the question was asked. Risks assigned higher frequency estimates were typically judged more likely when paired with risks with lower frequency estimates. Ratios of the direct estimates were similar to directly estimated ratios. Thus, these people seemed to have an internal 'scale' of relative risk, which they expressed consistently even with these unfamiliar tasks.

2. Judgements of *absolute* risk, however, were affected by the procedure. People told that 50,000 people die annually from auto accidents gave estimates two to five times higher than did people told that 1000 die annually from electrocution. Thus, people seemed to have less feeling for absolute frequency, rendering them sensitive to implicit cues given by how questions are posed (Poulton 1989; Schwarz 1999; Fagerlin and Peters, 2011).

3. Absolute risk judgements were less dispersed than were the corresponding statistical estimates. Although the latter varied over six orders of magnitude, individuals' estimates typically ranged over three to four. That 'compression' could reflect another judgemental bias, called *anchoring*, whereby judgements are drawn toward an initial value that draws their attention. With these anchors (electrocution, motor vehicle accidents), people overestimated small frequencies and underestimated large ones. That pattern might change with other anchors. For example, a lower anchor (e.g. botulism) should reduce (or perhaps eliminate) the overestimation of small frequencies, while increasing the underestimation of large ones.

4. Relative and absolute risk judgements seemed to reflect availability bias. Some causes of death (e.g. flood, homicide, tornadoes) received higher estimates than did others with similar statistical frequency. Typically, there were causes that were disproportionately reported in the news media. When told about the possibility of availability bias, participants could not improve their judgements, consistent with the finding that tracking frequency is such an automatic process that people do not realize how observations shape their perceptions (e.g. Koriat 1993).

Thus, Lichtenstein et al. (1978) found some response patterns that were affected by the procedure that was used (e.g. absolute estimates) and some that were not (e.g. relative risk judgements). A century of psychophysics research (Poulton 1989) has identified many other procedural details that can affect quantitative judgements. Determining how much those details affect any specific judgement requires studies examining their relative impact in that context. How important that effect (or any bias) is depends on the decision. Shifting fatality estimates by a factor of two to five might tip some decisions, but not others.

Fischhoff and MacGregor (1983) provide another example of response mode effects. They asked about the chances of dying (in the United States) among people afflicted with various maladies (e.g. influenza), in four ways: (1) how many people die out of each 100,000 who get influenza; (2) how many people died out of the 80 million who caught influenza last year; (3) for each person who dies of influenza, how many have it and survive; (4) 800 people died of influenza last year, how many survived? As in Lichtenstein et al. (1978), relative risk judgements were consistent across response modes, while absolute estimates varied greatly (over one to two orders of magnitude). They also found that people liked format (3) much less than the others—and were much less able to remember statistics reported that way. That

format also produced the most discrepant estimates, identifying it as a poor way to elicit or communicate risks.

Evaluative standards

Risk judgements can be evaluated in terms of their *consistency* or their *accuracy*. Evaluating consistency requires asking logically related questions and comparing the answers (e.g. do risk estimates increase with increasing exposure?). Evaluating accuracy requires asking questions that are sufficiently precise to be compared to sound risk estimates (e.g. Chapter 7.5). Without sound scientific estimates, individuals' judgements may be compared to a standard that they would reject. For example, after the 9/11 attacks in New York by terrorist-commandeered aeroplanes, some observers claimed that some Americans had increased their risk level by driving, rather than flying. These claims were based on historical risk statistics. However, it was difficult to ascertain the safety of aviation at that time as the US aircraft fleet was grounded, whereas the historical statistics used for driving encompassed all drivers, including the young, elderly, and drinkers, and were not specific to those drivers who had changed their transportation modes. Even if these historical statistics were valid, other factors must have affected the drivers' decisions such as the cost and hassle of flying during that period. As a general rule, one cannot infer risk judgements from risk decisions without knowing the other factors involved.

Probability judgements

The sensitivity of quantitative judgements to methodological details might suggest avoiding them in favour of verbal quantifiers (e.g. likely, rare). Indeed, some researchers hesitate to elicit probabilities at all, fearing that the questions will exceed laypeople's cognitive capabilities. That hesitation is strengthened by evidence of lay innumeracy (Fagerlin and Peters 2011). However, even imperfect measures can have value, if their strengths and weaknesses are understood. The research literature on eliciting probability judgements is vast (O'Hagan et al. 2006). Findings relevant to public health researchers and practitioners include:

1. People often prefer to provide verbal judgements and receive numeric ones, given that numeric responses require more effort and incur greater accountability (Erev and Cohen 1990).
2. Verbal quantifiers are often interpreted differently across people and situations (e.g. rare disease vs. rare sunny day), making it hard to know what those terms mean, in situations without established usage norms (Budescu and Wallsten 1995; Schwarz 1999).
3. People can use well-designed numeric scales as well as verbal ones. For example, Woloshin et al. (1998) found similar performance and satisfaction with linear and log-linear probability scales as with verbal ones.
4. Numeric probability judgements often have good *construct validity*, in the sense of correlating sensibly with other variables. For example, Fischhoff et al. (2000) found that teens who gave higher probabilities of becoming pregnant also reported more sexual activity; teens giving higher probabilities of getting arrested also reported more violent neighbourhoods.
5. Misinformation and mistaken inferences can bias probability judgements, as when one's own care in driving is more available that that of other drivers, making one feel safer than average.

6. Probability judgements can be deliberately biased, when people respond strategically. For example, Christensen-Szalanski and Bushyhead (1993) found that physicians overestimated the probability of pneumonia, fearing that unlikely cases might be neglected. Weather forecasters may overstate the probability of precipitation, in order to keep people from being caught unprotected (Lichtenstein et al. 1982).

7. Transient emotions can affect judgements. For example, anger increases optimism, fear the opposite (Lerner and Keltner 2001), with effects large enough to tip close decisions.

8. Judgements of the probability of being correct are moderately correlated with how much people actually know. For example, Fischhoff et al. (1977) had people choose the larger of two causes of death (from Lichtenstein et al. 1978), and then give the probability of having chosen correctly. In relative terms, people were correct more often when they were more confident. In absolute terms, *overconfidence* (e.g. being 90 per cent confident with 75 per cent correct choices) is typical with hard tasks, underconfidence with easy ones.

9. Probability judgements can vary by response mode (e.g. odds vs probabilities, probabilities vs relative frequencies, judgements of individual or grouped items) (Griffin et al. 2003).

10. Some numeric values are treated specially. For example, people seldom use fractional values; when uncertain what to say, people sometimes say 50 in the sense of 50–50, rather than a numeric probability (Bruine de Bruin et al. 2000).

11. Probability judgement processes mature by middle adolescence. For example, teens are no more likely than adults to believe in their own adolescent invulnerability (Quadrel et al. 1993); indeed, unlike adults, many teens greatly exaggerate their probability of premature death (Fischhoff et al. 2000).

12. People differ in their ability to use probabilities, with lower ability correlated with poorer performance on other tasks and with life outcomes that require decision-making competence (Bruine de Bruin et al. 2007b).

13. The use of probabilities can sometimes be improved with even a single round of prompt, intense feedback (Lichtenstein and Fischhoff 1980).

14. Experts' judgements are often imperfect, when forced to go beyond established knowledge and calculations (O'Hagan et al. 2006).

A test of any measure is its predictive validity. Even though risk decisions often involve choices among options with non-risk outcomes (which might outweigh risk concerns), Brewer et al. (2007) found that risk judgements alone have predictive value. Similarly, teens' probability judgements predict major events in their lives (e.g. pregnancy, incarceration), one to 5 years hence (Fischhoff 2008). Pointing to probability judgements that are higher than actual risks, some researchers have argued that public health communications have worked too well, producing exaggerated fears of smoking (Viscusi 1992) and breast cancer (Black et al. 1995).

Defining risk

Studies like Lichtenstein et al. (1978) measure 'risk' perceptions, if 'risk' means 'chance of death'. However, even among experts, 'risk' has multiple meanings (Fischhoff et al. 1984; National Research Council 1996). 'Risk' might mean just death or it might also include other outcomes, such as morbidity and trauma. Even if 'risk' only considers fatalities, it might be

measured in terms of probability of death, expected life years lost, or deaths per person exposed (or per hour of exposure). Each definition entails an ethical position. For example, *probability of death* treats all deaths (and lives) equally, whereas *life-years lost* places extra weight on deaths of young people and from injury (e.g. drowning, driving, workplace hazards), as each incurs many lost years, compared to deaths from chronic illnesses. Adding morbidity and trauma would heighten concern for alcohol and illegal drugs, which can ruin lives without ending them.

Without clear, shared definitions, people can unwittingly speak at cross purposes, when addressing 'risks'. Clarifying definitions has long been central to risk research. Before considering that research, it is worth noting that 'risky' (or 'safe') is sometimes used as a discrete variable, treating activities as risky (or safe) or not. Such shorthand says little, without defining the threshold of concern. Calls for 'safe' products can be unfairly ridiculed, if the demand for reasonable risk is equated with zero risk. Such demands are seen in the various *precautionary principles*, identifying risks seen as too great to countenance (DeKay et al. 2002). However, even those calls may be more about uncertainty than risk, reflecting aversion to hazards that science does not understand (Löfstedt et al. 2002).

Catastrophic potential

One early risk perception study asked experts and laypeople to estimate the 'risk of death' from 30 activities and technologies (Slovic et al. 1979). These judgements correlated more strongly with statistical estimates of average-year fatalities for experts than they did for laypeople. However, when asked to estimate 'fatalities in an average year', experts and laypeople responded similarly. Comparing the two sets of judgements suggested that lay respondents interpreted 'risk of death' as including catastrophic potential, reflecting the expected deaths in non-average years. If so, then experts and laypeople agreed about the risk of routine (average year) deaths (for which the science is often good), but disagreed about possible anomalies (for which the science is naturally much weaker). Thus, when experts and laypeople disagree about risks, they might be seeing the facts differently or they might be looking at different facts, ones relevant to their definition of 'risk' (National Research Council 1989). People might consider catastrophic potential because they care more about lives lost at once than lost individually or because catastrophic potential suggests hazards that might spin out of control (Slovic et al. 1984).

Dimensions of risk

Beginning with Starr (1969), many features, like uncertainty and catastrophic potential, have been suggested as affecting definitions of risk (Lowrance 1976). In order to reduce that set to a manageable size, Fischhoff et al. (1978) asked members of a liberal civic organization to rate 30 hazards on nine such features. Factor analysis on mean ratings identified two *dimensions*, which accounted for 78 per cent of the variance. Fig. **7.6.2** plots factor scores in this 'risk space'. Similar patterns emerged with students, members of a conservative civic organization, and risk experts, suggesting that people think similarly about such factors, even when they disagree about how specific hazards stack up.
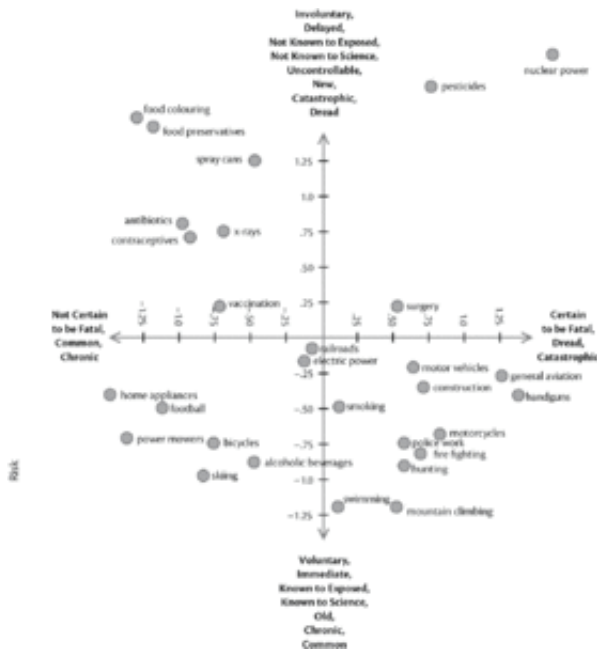
Fig. 7.6.2
Location of 30 hazards within the two-factor space obtained from members of a civic group who rated each activity or technology on each of nine features. Ratings were subjected to principal components factor analysis, with a varimax rotation.

Reproduced from Fischhoff, B. and Kadvany, J., *Risk: A Very Short Introduction,* Figure 4, p. 30, Oxford University Press, Oxford, UK, Copyright © Baruch Fischhoff and John Kadvany 2011, by permission of Oxford University Press. Previously adapted from Fischhoff, B. et al., How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits, *Policy Sciences,* Volume 9, Issue 2, pp. 127–152, Copyright © 1978. Reproduced with kind permission from Springer Science and Business Media.

Hazards high on the vertical factor (e.g. food colouring, pesticides) were rated as new, unknown, and involuntary, with delayed effects. Hazards high on the horizontal factor (e.g. nuclear power, commercial aviation) were rated as fatal to many people, if things go wrong. The factors were labelled *unknown* and *dread*, respectively, and might be seen as capturing cognitive and emotional aspects of people's concern.

Many studies following this 'psychometric paradigm' have found roughly similar dimensions, using differing elicitation modes, scaling techniques, items, and participants (Slovic 2001). When a third dimension emerges, it typically reflects the scope of the threat, labelled *catastrophic potential*. The position of hazards in the space correlates with attitudes towards them, such as how stringently they should be regulated. Analyses of mean responses, as in the figure, are best suited to predicting aggregate (societal) responses. Individual differences have also been studied (e.g. Vlek and Stallen 1981; Arabie and Maschmeyer 1988).

Risk comparisons

The multidimensionality of risk means that hazards similar on some dimensions can still evoke

quite different responses. This fact is neglected in appeals to accept one risk because one has accepted another risk with some similarities (Fischhoff et al. 1984). A common kind of such 'risk comparison' presents the statistical risks from many hazards in common terms (e.g. arguing that people who eat peanut butter should accept nuclear power because both a tablespoonful of peanut butter and 50 years living by a nuclear power plant create a one-in-a-million risk of premature death). (For a summary of the problems with such risk comparisons, see National Research Council (2006).)

One way to improve the legitimacy of risk comparisons is to involve users in setting them. The US Environmental Protection Agency (1993) followed this strategy in facilitating dozens of regional, state, and national 'risk-ranking exercises', in which participants identified the dimensions important to them, then deliberated priorities, supported by technical staff providing relevant evidence. Letting participants choose the dimensions made their exercise more relevant, but reduced comparability across exercises. Florig et al. (2001) developed a method for standardizing such comparisons, based on the risk dimensions research (Table 7.6.1). The UK government has endorsed a variant (HM Treasury 2005).

Table 7.6.1 A standard multidimensional representation of risks			
Number of people affected	**Degree of environmental impact**	**Knowledge**	**Dread**
Annual expected number of fatalities: 0–**450**–600 (10% chance of zero)	Area affected by ecosystem stress or change **50** km^2	Degree to which impacts are delayed **1**–**10** years	Catastrophic potential**1000** times expected annual fatalities
Annual expected number of person-years lost: 0–**9000**–18,000 (10% chance of zero)	Magnitude of environmental impact **Modest** (15% chance of large)	Quality of scientific understanding **Medium**	Outcome equity **Medium** (ratio = 6)

Source: data from Willis, H.H. et al., Aggregate and disaggregate analyses of ecological risk perceptions, *Risk Analysis*, Volume 25, Issue 2, pp. 405–428, Copyright © 2005.

Qualitative assessment

Event definitions

Once adequately defined, 'risk' can be estimated. For risk assessors, that means specifying such details as the frequency and timing of intercourse, contraceptives used, and partners' physical state—when estimating the risk of pregnancy. Two experts with different definitions may see the same data and produce different estimates. So may laypeople asked for their

perceptions of risk, but forced to guess at what exactly is meant. Consider this question from a prominent national survey: 'How likely do you think it is that a person will get the AIDS virus from sharing plates, forks, or glasses with someone who has AIDS?' After answering this question, US college students were asked what they had inferred about the kind and amount of sharing. Most agreed about the kind, with 82 per cent selecting 'sharing during a meal' from a set of options. However, they disagreed about the frequency, with 39 per cent selecting 'a single occasion', 20 per cent 'several occasions', 28 per cent 'routinely', and 12 per cent uncertain (Fischhoff 1996). Respondents making different assumptions were, in effect, answering different questions, whose meaning researchers must guess, if they are to offer any conclusions about lay risk perceptions.

Laypeople are, similarly, left guessing when experts communicate about risks ambiguously (Fischhoff 1994). For example, McIntyre and West (1992) found that teens knew that 'safe sex' was important, but disagreed about what it entailed. Downs et al. (2004b) found that teens interpret 'it can only take once' as meaning that they will get pregnant after having sex once. If they do not, some infer that they are infertile, encouraging unsafe sex. Murphy et al. (1980) found people divided over whether '70 per cent chance of rain' referred to: (1) the area receiving rain, (2) the time it would rain, (3) the chance of some rain anywhere, or (4) the chance of some rain at the weather station (the correct answer). Fischhoff (2005a) describes procedures for making sure that experts and laypeople are talking about the same thing, when they communicate about risks.

Supplying details

The details that people infer, when given ambiguous and incomplete risk questions or messages, reveal their intuitive theories. For example, teens who thought aloud while judging the probabilities of ambiguous events (like that about sharing plates, etc., with someone with AIDS), noticed many unstated details, including ones that would affect scientific risk estimates (Fischhoff 1994). For example, they wondered about the 'dose' of most risks (e.g. the amount of drinking and driving, when judging the probability of an accident), when it was missing from a question. An exception was not thinking about the amount of sex involved, when judging the risks of pregnancy and HIV transmission. Teens seemed to believe that an individual is either vulnerable or not, making the number of exposures immaterial. Sometimes they considered variables unrelated to risk, such as how well partners know one another. In order to dispel such misunderstanding, Downs et al. (2004a) explicitly addressed how partners could fail to self-diagnose sexually transmitted infections (STIs)—in an interactive DVD that successfully reduced adolescent sexual risks.

Cumulative risk—a case in point

There is no full substitute for directly studying the beliefs that people bring to and take away from risk messages, especially when recipients come from cultures and social circumstances different than those of the communicators. However, the research literature provides a basis for anticipating those beliefs (Fischhoff et al. 2011). For example, optimism bias is so widespread that one can assume that people see themselves as facing less risk than other people, whenever some personal control seems feasible. Similarly, teens' insensitivity to the amount of sex, when judging STI risks, reflects a well-known insensitivity to how risks accumulate over repeated exposures. Thus, people cannot be expected to infer the

cumulative accident risk from repeatedly driving without a seat belt (Slovic et al. 1978) or the pregnancy risk from having sex without effective contraceptives (Shaklee and Fischhoff 1990). One corollary of this insensitivity is not realizing the cumulative impact of small differences in single-exposure risks (e.g. slightly better contraceptives, wearing a seat belt). People similarly underestimate exponential growth (e.g. Wagenaar and Sagaria 1975; Frederick 2005).

For example, Linville et al. (1993) had college students judge the probability of transmission from an HIV-positive man to a woman from 1, 10, or 100 cases of protected sex. For one case, the students' median estimate was 0.10, much higher than then-current public health estimates —despite using a log-linear response mode that facilitated expressing very low probabilities (Woloshin et al. 1998). The median estimate for 100 contacts was 0.25, a more accurate estimate, but much too small given their one-case estimates. Given the inconsistency in these beliefs, researchers studying risk perceptions must ask about both, in order to get a full picture, and educators seeking to inform risk beliefs need to communicate them both, in order to create a full picture.

Mental models of risk processes

The role of mental models

As mentioned, when people lack explicit information about the magnitude of a risk (or benefit), they must infer it. Judgemental heuristics, like availability, provide one class of inferential rules for deriving specific estimates from general knowledge. A second class of inferential rules draws on individuals' *mental models* of the general processes that create and control risks in order to estimates those risks, follow discussions about them, and generate choice options. The term 'mental model' refers to the intuitive theories supporting such inferences. Mental models have a long history in psychology, having been studied for topics as diverse as how people understand physical processes, international tensions, complex equipment, energy conservation, climate change, interpersonal relations, and drug effects (Meyer et al. 1985; Ericsson and Simon 1993; Sterman and Sweeney 2002).

However sound these inferences, they can produce erroneous conclusions when mental models contain flawed assumptions (or 'bugs'). For example, not realizing how quickly the risks of pregnancy and STIs accumulate over sex acts could make other knowledge seem irrelevant. Bostrom et al. (1992) found that many people knew that radon was a colourless, odourless, radioactive gas, but overestimated its risks because they also thought that radioactivity meant permanent contamination. However, radon's by-products (or 'progeny') have short half-lives, meaning that once intrusion of the gas stops, the problem disappears. However, while it persists, rapid decay means rapid energy release. Homeowners unaware of these facts might reasonably decide not to test for radon—the problem doesn't seem urgent and there is nothing to do anyway if they find a problem.

Morgan et al. (2001) offer a general approach appropriate to studying mental models for complex, uncertain processes, like those of many public health risks. The approach begins by creating a formal (or 'expert') model, summarizing relevant scientific knowledge, with enough conceptual precision to allow computing quantitative predictions, were its data needs met (Fischhoff et al. 2006). A common formalism is the influence diagram (Howard 1989). Fig. **7.6.3** shows such a diagram for radon. An arrow means that the value of the variable at its head depends on the value of the variable at its tail. Thus, the lungs' particle clearance rate

depends on individuals' smoking history. Other examples include STIs (Fischhoff et al. 1998), breast implants (Byram et al. 2001), sexual assault (Fischhoff 1992), Lyme disease, falls, sexual assault, breast cancer, vaccination, infectious disease, and nuclear energy sources in space (Morgan et al. 2001; Fischhoff 2005b; Downs et al. 2008).
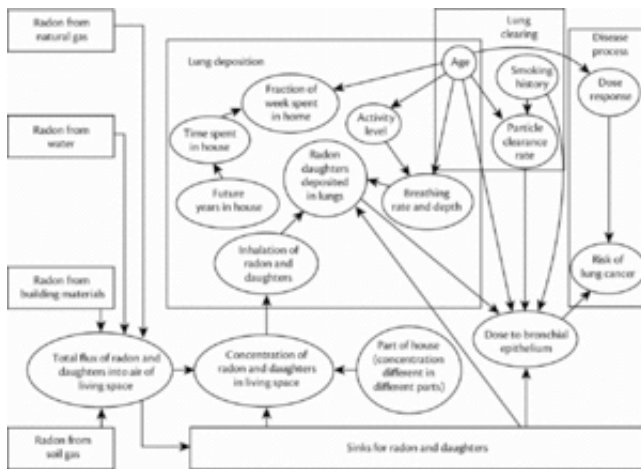

Fig. 7.6.3
Expert influence diagram for health effects of radon (in a home with a crawl space). This diagram was used as a standard and as an organizing device to characterize the content of lay mental models.

Reprinted with permission from Morgan, M.G. et al., Communicating Risk to the Public: First, Learn what people know and believe, *Environmental Science and Technology*, Volume 26, pp. 2048–56, Copyright © 1992, American Chemical Society.

The research continues with open-ended one-on-one interviews, structured around the model, eliciting lay beliefs in their intuitive formulation. Those 'mental model' interviews begin with general questions, asking respondents what they believe about the topic, then to elaborate on each issue raised. The interviews are non-judgemental, seeking to understand, not evaluate respondents' perspectives. After exhausting responses to general questions, interviewers ask increasingly pointed ones, starting with general processes (e.g. exposure, effects, mitigation), and proceeding to specific issues (e.g. 'How does the amount of sex (or number of partners) affect HIV risk?'; 'What does 'safe sex' mean?'). A variant has people think aloud while sorting photographs by their relevance, hoping for insights into topics that were otherwise missed. For example, seeing a supermarket produce section prompted some respondents to say that radon might contaminate plants (Bostrom et al. 1992).

Once transcribed, interviews are coded into the formal model, adding new elements raised by respondents, marked as either misunderstandings or expertise (e.g. knowledge about how equipment really works). The precision of the formal model typically allows reliable coding. Once mapped, lay beliefs can be analysed in terms of their accuracy, relevance, specificity, and focus. Coding for accuracy can reveal beliefs that are correct and relevant, wrong, vague, peripheral, or general (e.g. radon is a gas). For example, Bostrom et al. (1992) found that most respondents, drawn from civic groups, knew that radon is a gas (88 per cent), which concentrates indoors (92 per cent), is detectable with a test kit (96 per cent), comes from

underground (83 per cent), and can cause cancer (63 per cent). However, many also believed erroneously that radon affects plants (58 per cent), contaminates blood (38 per cent), and causes breast cancer (29 per cent). Few (8 per cent) mentioned that radon decays. The interviews led to a structured survey suited to assessing the prevalence of beliefs in larger samples, with questions having ecological validity, in the sense of sampling the key topics in the formal model (Bruine de Bruin et al. 2007a).

From risk beliefs to risk decisions

The adequacy of risk perceptions depends on the decisions that depend on them. Some decisions require precise estimates, others just a rough idea. For example, von Winterfeldt and Edwards (1986) showed that many decisions with continuous options (e.g. invest US$X) are insensitive to the precise values assigned to the probabilities and utilities of possible outcomes. Dawes et al. (1989) showed that choices with discrete options (e.g. choosing graduate candidates) are often insensitive to exactly how predictors or outcomes are weighted, meaning that simple linear (weighted-sum) models may do as well as more complicated ones. Thus, any model that considers the probability and magnitude of consequences should have some success in predicting behaviour, if researchers have some idea about the topics on decision-makers' minds. On the other hand, because many such models will do reasonably well, they provide little insight regarding the underlying processes.

Feather (1992) provides a general account of such *expectancy- value* (probability-consequence) models, which predict decisions by multiplying ratings of the likelihood and desirability of potentially relevant consequences. The health-belief model and the theory of reasoned action fall into this general category. For example, Bauman (1980) had seventh graders rate the importance, likelihood, and valence (positive or negative) of 54 possible consequences of using marijuana. A 'utility structure index', computed from these three judgements, predicted about 20 per cent of the variance in respondents' reported marijuana usage.

The template for studying these perceptions is a *decision tree* with the options, relevant outcomes, and uncertain events linking the two. Fig. **7.6.4** shows a simple decision tree, for men considering the dietary supplement, saw palmetto, for symptomatic relief of benign prostatic hyperplasia. The choice (the square node on the left) leads to a sequence of events (the circular *uncertain event* nodes), resulting in the outcomes (or consequences) on the right. The success of a structured model (e.g. Bauman 1980) depends on how well it captures the issues that occupy decision-makers. In identifying those elements, researchers can draw on previous research, convention, or intuitions—or by eliciting them from decision-makers. The greater the social distance between the experts and the decision-makers, the more important such elicitation becomes—lest experts miss options, uncertainties, or outcomes that occupy decision-makers, but would never occur to them, or vice versa.
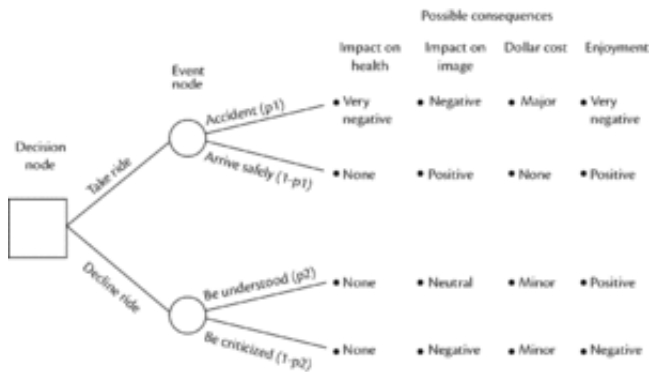
Fig. 7.6.4
A simple decision tree for whether to take saw palmetto for benign prostatic hyperplasia.

Reproduced From Fischhoff, B. and Quadrel, M.J. Adolescent alcohol decisions, *Alcohol Health and Research World*, Volume 15, pp. 43–51, 1991.

Effective elicitation typically requires prompting different ways of looking at a decision, so that respondents do not get locked into a narrower perspective than would occur in life (Schwarz 1999). For example, Beyth-Marom et al. (1993) had teens work out possible consequences of either accepting a risky option (e.g. drinking and driving, smoking marijuana) or rejecting it. Although accepting and rejecting are formally complementary actions, they can stimulate different thought processes. In this study, participants who thought about accepting risky options produced more consequences (suggesting that action is more evocative than inaction), a higher ratio of bad to good consequences (suggesting that risks are more available from that perspective), and fewer references to social consequences (suggesting that social pressure is more salient when resisting temptation than when yielding to it). When participants thought about making choices repeatedly, rather than just once, they often produced different consequences (e.g. repeatedly 'accepting an offer to smoke marijuana at a party' evoked more mentions of social reactions than did thinking about doing it once). Parents of these teens cited similar possible outcomes, except for being more likely to mention long-term consequences (e.g. ruining career prospects). From this perspective, if parents and teens see the choices differently, it is not because they see different outcomes as possible, but because they disagree about how likely and important those outcomes are. These different perspectives would be hidden with structured surveys that elicit ratings of fixed, predetermined consequences.

Fischhoff (1996) reports a study imposing even less structure, with 105 teens asked to describe three difficult personal decisions in their own words. These descriptions were coded in terms of their content (which choices trouble teens) and structure (how they were formulated). For example, none of the teens mentioned a choice about drinking-and-driving, while many described drinking decisions. Few of their decisions had option structures as complicated as Fig. **7.6.4**. Rather, most had but one option (e.g. whether to attend a party with drinking). Judging by Beyth-Marom et al.'s (1993) results, teens looking at that option saw different decisions than did teens focusing on other possible options (e.g. going somewhere else) or multiple options. Experimental research has found that the opportunity costs (foregone benefits) of neglected options are less visible than are their direct consequences (Thaler

1991). For example, the direct risks of vaccinating children can loom disproportionately larger than the indirect risks of not vaccinating them (Ritov and Baron 1990).

Different methods for eliciting decision-makers' perspective have different, often complementary strengths and weaknesses (Ericsson and Simon 1993). Structured methods (e.g. surveys) can omit important aspects of decisions or express them in unfamiliar terms. Open-ended methods (e.g. mental models interviews) allow people to say whatever is on their minds in their own terms, but require tight control lest researchers influence what is said. Combining methods can provide a rounded picture, especially when a formal analysis ensures their comprehensiveness. Unlike commercial research, scientific studies rarely use focus groups, except for the initial generation of ideas. Indeed, the inventor of focus groups, Robert Merton (1987) rejected them as sources of evidence, given the unnatural discourse of even the best-moderated group, the difficulty of hearing individuals out, and the impressionistic coding of contributions. He preferred *focused interviews*, akin to mental models interviews without the normative analysis. Whichever methods researchers use, they are likely to miss the mark unless they listen to decision-makers' perspectives, before imposing structured methods or designing communications.

Creating communications

Selecting information

Communication design begins by selecting content. The gold standard is a normative analysis, identifying the information most relevant to the choices that the communication is meant to inform. In practice, though, the content-selection process often is ad hoc, with experts intuiting 'what people ought to know' (Nickerson 1999). Poorly selected information can waste recipients' time, take the place of relevant content, or bury facts that people need to know among others that might only be nice to know. Poorly selected information can erode recipients' faith in the experts responsible for communications (and in the institutions employing them), by showing insensitivity to their informational needs ('Why are you telling me X, when I need to know Y?'). It can also undermine experts' faith in their audience, if they fail to realize that their messages have missed the mark. For example, Florig and Fischhoff (2007) found that it was impractical for many individuals to secure and store items on official lists of emergency provisions. Recipients of such advice might ask why they were being asked to do the impossible (Fischhoff, 2011).

The logic of setting information priorities is straightforward: begin with the facts that will have the greatest impact, if they are properly understood. In economics terms, that means creating a 'supply curve' for facts, ordered by their importance. That task can be formalized in 'value of information analysis' (vonWinterfeldt and Edwards 1986; Sox et al. 2007), as used by Merz et al. (1993) in setting priorities when securing informed consent for medical procedures, with carotid endarterectomy as a case study. Scraping out the main artery to the brain can reduce stroke risk, but also cause many problems, including strokes. Attempting to communicate all these risks could easily overwhelm patients. The research identified the risks that mattered most by creating a population of hypothetical patients, varying in their physical condition and health preferences, all of whom would want the procedure were there no side effects (and were money no object). The analysis then asked what percentage of these patients should decide against the surgery, upon learning about each possible side effect. It found that only

three of the many side effects (death, stroke, facial paralysis) were sufficiently likely and severe to change many decisions. Although nothing should be hidden, communications should be sure to get the few key facts across.

At times, people are not required to make a specific choice, but are just trying to understand a situation that could pose many decisions (e.g. a newly diagnosed disease, food-borne illness). The same logic of prioritization applies here as well. Communications should focus on the information that is most useful for predicting the outcomes that matter most (e.g. the critical signs of health problems, the key determinants of food safety). That information completes the *mental model* that people need to monitor their environment, generate action options, and follow discussions on the topic (Morgan et al. 2001). Here, too, building on individuals' existing knowledge allows focusing communications on critical gaps (while also demonstrating that the experts know what their audience already knows). For example, Downs et al. (2004a) found that most teens knew so much about HIV/AIDS that communications could focus on a few critical gaps, such as how risks mount through repeated exposure and how hard it is for sexual partners to self-diagnose their own disease status.

An essential part of the content of any communication is the strength of the evidence supporting it (O'Hagan et al. 2006; Politi et al. 2007). The most dangerous beliefs are those held with too great or too little confidence, leading to overly risky or overly cautious actions. Campbell (2011) shows ways to represent uncertainty graphically. Schwartz and Woloshin (2011) showed how much can be conveyed with text describing the quality of the data (e.g. the length, size, and quality of clinical trials). Funtowicz and Ravetz (1990) showed how to characterize the quality of the underlying science, including its *pedigree* (e.g. the extent to which empirical patterns are supported by theory). As an example of how an assessment of uncertainty can inform choices, a meta-analysis (Fortney 1988) concluded, with great confidence, that oral contraceptives may increase a non-smoking woman's life expectancy by up to 4 days and decrease it by up to 80 days. Moreover, the research base was so large that no conceivable study could materially change those bounds.

Formatting information

Once selected, information must be presented. Reimer and Van Nevel (1999) and Wogalter (2006) provide important pointers on research on alternative displays. They note, for example, that comprehension improves when: (1) text has a clear structure, corresponding to recipients' intuitive representation; (2) there is a clear hierarchy of information; and (3) readers receive *adjunct aids*, such as highlighting, advanced organizers (showing what to expect), and summaries.

Scientifically established design principles provide a point of departure for arranging information. These are better 'best guesses' than those informed merely by intuition. Their success in any specific application is an empirical question, though, which can be studied with standard usability testing procedures, such as seeing how long it takes users to find designated pieces of information, how often they reach the wrong information, and how likely they are to realize that (Wogalter 2006). Riley et al. (2001) developed a general method for evaluating the adequacy of communications, drawing on basic research into search patterns. Taking methylene chloride-based paint stripper as a case study, the method begins by identifying critical information (in this case the steps that most effectively reduce exposures to

the chemical and its by-products). It then evaluates product labels by seeing what risk-related information would be found by users who search in different ways. For example, a label might reveal critical information to someone who reads the first five items, but not someone who only reads the instructions or just highlighted material. Actual experience will depend on the prevalence of these search patterns (e.g. what percentage of users look at black box warnings or have instructions read to them). Unless the communication format fits users' natural search patterns, its information might be hidden in plain sight. Riley et al. found that some paint stripper products made critical, useful precautionary information accessible to any reader, while some helped only some readers (e.g. those who read warnings first), and some omit critical information altogether.

Evaluating communications

However sound their theoretical foundations, communications must be empirically evaluated (National Research Council 1989; Slovic 2001). One should no more release an untested health communication than an untested drug. Indeed, communications are part of any medical product or procedure, shaping when it is chosen, how it is used, and whether problems are noticed in time to be remedied and reported. Arguably, evidence about the effectiveness of such communications should be part of the evidence submitted when requesting approval of a product, or when conducting post-licensing surveillance of its benefits and risks in actual use.

A communication is *adequate* if it:

- ◆ Includes the information that recipients need, in order to make decisions about risks.
- ◆ Affords them access to that information, given their normal search patterns.
- ◆ Allows them to comprehend that information, with a reasonable effort.

Applying each of these three tests requires evidence. Knowing what information people need requires learning their goals, which may differ from those of the experts providing the information. Knowing whether people can find the information that is there requires observing how they search. Knowing how much they comprehend requires seeing how well they have mastered the content.

As seen in the references to this chapter, applying these tests to a publication standard is a serious undertaking, requiring professional training. However, simple versions of each test are within the reach of any communicator. The US Food and Drug Administration's *Communicating Risks and Benefits: An Evidence-Based User's Guide* (Fischhoff et al. 2011) ends each chapter with a section on how to conduct evaluations at no cost, a small cost, and a cost commensurate with the stakes riding on effective communication. Central to all forms of evaluation is listening, without presuming to know recipients' goals, beliefs, uncertainties, emotions, or modes of expression. In order to identify individuals' information needs, ask how they see the risks in the context of their lives. In order to see how easily people can access information, watch as they search for it in existing sources (e.g. online) and drafts of proposed communications (Downs et al. 2008). In order to assess a communication's comprehensibility, ask people to recall it, paraphrase it, make inferences from it, or create scenarios using it (Bruine de Bruin et al. 2009).

These are all structured ways of conducting conversations about technical topics of mutual

interest, designed to bridge some of the social distance between experts and lay people. For these methods to succeed, they also need to bridge any perceived status difference. Thus, they must be framed as testing the communications, not the recipients, in order to help experts to help the public. Almost any open-minded data collection is better than none. Thus, even a few open-ended, one-on-one interviews might catch incomprehensible or offensive material. The core presumption of risk communication should be that, if lay people have not learned facts that matter to them, the expert community must have failed to get that information across to them. Only if scientific resources have been exhausted should it be assumed that laypeople are incapable of learning the required information. The stakes riding on facilitating lay decision-making should justify that investment and humility. Amateurish, unscientific communications can be worse than nothing, by holding audience members responsible for failing to understand risks when the information was missing, inaccessible, or incomprehensible.

The science of communication can guide both *persuasive communications*, designed to influence individuals to act in ways determined by the communicator, and *non-persuasive communications*, designed to help individuals identify actions in their own best interest. The two kinds of communication converge when persuasive communicators establish that they are influencing people in ways that they would accept as being 'for their own good' (Thaler and Sunstein 2009). Without studying people's goals, however, one risks imposing experts' views on them. For example, in a study mentioned earlier, Bostrom et al. (1992) found people who rejected persuasive communications that advocated testing for radon because they wanted to avoid creating evidence that could complicate selling their homes. Fischhoff (1992) reports on the conflicting advice given to women about reducing the risk of sexual assault, reflecting differences in the goals that experts attribute to the women (and in beliefs regarding the effectiveness of self-defence strategies) (Farris and Fischhoff, 2012). Slovic and Fischhoff (1983) describe how reasonable individuals may 'defeat' safety measures by gaining more benefit from a product (e.g. driving faster with a car that handles better), frustrating policymakers concerned solely with safety.

Managing communication processes

In order to communicate effectively, organizations require four kinds of expertise:

> **1**. *Subject matter specialists*, who can identify the processes that create and control risks (and benefits).
> **2**. *Risk and decision analysts*, who can estimate the risks (and benefits) most pertinent to decision-makers (based on subject matter specialists' knowledge).
> **3**. *Behavioural scientists*, who can assess decision-makers' beliefs and goals, guide the formulation of communications, and evaluate their success.
> **4**. *Communication practitioners*, who can create communication products and manage communication channels, getting messages to audiences and obtaining feedback from them.

The work of these experts must be coordinated, so that they play appropriate roles. For example, behavioural scientists should not revise text (trying to improve its comprehensibility) without having subject matter specialists check that the content is still accurate; subject matter specialists should not slant the facts according to their pet theories of how the public needs to

be alarmed or calmed. Without qualified experts, these roles will be filled by amateurs, imperilling the organization and its public.

Conclusion

Effective risk communication is essential to managing risks in socially acceptable ways. Without it, individuals are denied the best possible chances of making sound choices—before, during, and after problems arise. As a result, they may suffer avoidable injury, along with the insult of feeling that the authorities have let them down, by not creating and disseminating the information that they needed, in a timely, comprehensible way. One should no more expose individuals to an untested risk communication than to an untested medical product or procedure.

Effective risk communication focuses on the decisions that people face. Without that focus, one cannot know what information they need. Sound risk management requires not only communicating that information, but also creating it, both through risk analyses, summarizing existing research (see Chapter 7.5), and new research creating the basis for risk analyses (most other chapters in this textbook). As a result, effective risk communication cannot be just an afterthought, letting the public know what the authorities have decided. Rather, it must be central to risk management, as part of disciplined, continuing, two-way communication between decision-makers and the authorities.

This chapter has focused on measurement, rather than on general theories about how people perceive and respond to risks. That is because critical details vary across risk decisions and decision-makers. Sweeping generalizations about what 'people do' or 'people think' or 'people want' undermine the attention to detail that responsive risk communications require. Separate research programmes could be dedicated to communicating the science presented in many chapters in the textbook, ensuring that the public gets full value from that science. However, the methods for studying judgement and decision-making are sufficiently general and well understood that they could be applied in any domain, and for any form of information dissemination. Given a well-characterized decision or risk, it is relatively straightforward, if technically demanding, to assess lay (or expert) perceptions. If decision-makers' risk (and benefit) perceptions have been measured well, their choices can often be roughly predicted with simple linear models (Dawes et al. 1989). More precise prediction requires more detailed understanding of the cognitive processes shaping these beliefs, as well as an understanding of the emotional, social, economic, and other factors impinging on specific decisions. Prediction may not be that important, when the public health goal is helping people to make the best choices or empowering them to change their circumstances.

Meeting the challenge of effective risk communication requires coordinating the activities of four kinds of experts: subject matter specialists, risk and decision analysts, behavioural scientists, and communication practitioners. Assembling those teams requires leadership, seeing communication as being essential to the public health mission. The research itself is inexpensive, relative to the stakes riding on sound risk decision-making, both for individuals and for the public health organizations expected to serve them. There is no good reason for the measurement of risk perceptions and the evaluation of risk communications to use less than the readily available methods described here. There is no good reason to ignore well-established results, such as the multidimensional character of 'risk', the problems with verbal

quantifiers, and the need to help people to understand how risks mount up through repeated exposure. Ad hoc communications might reflect sound intuition, but they deserve less trust than scientifically developed ones.

By definition, better risk communication should help its recipients to make better choices. It need not make the communicators' lives easier—recipients may discover bona fide disagreements with the communicators and their institutions. What it should do is avoid conflicts due to misunderstanding, increasing the light-to-heat ratio in risk management, leading to fewer but better conflicts (Fischhoff 1995).

Acknowledgement

References

Arabie, P. and Maschmeyer, C. (1988). Some current models for the perception and judgment of risk. *Organizational Behavior and Human Decision Processes*, *41*, 300–29.

Bauman, K.E. (1980). *Predicting Adolescent Drug Use: Utility Structure and Marijuana*. New York: Praeger.

Beyth-Marom, R., Austin, L., Fischhoff, B., *et al*. (1993). Perceived consequences of risky behaviors. *Developmental Psychology*, *29*, 549–63.

Black, W.C., Nease, R.F., and Tosteson, A.N.A. (1995). Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. *Journal of the National Cancer Institute*, *8*, 720–31.

Bostrom, A., Fischhoff, B., and Morgan, M.G. (1992). Characterizing mental models of hazardous processes: a methodology and an application to radon. *Journal of Social Issues*, *48*(4), 85–100.

Breakwell, G.M. (2007). *The Psychology of Risk*. Cambridge: Cambridge University Press.

Brewer, N.T., Chapman, G.B., Gibbons, F.X., *et al*. (2007). Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. *Health Psychology*, *26*, 136–45.

Bruine de Bruin, W., Downs, J.S., Fischhoff, B., and Palmgren, C. (2007a). Development and evaluation of an HIV/AIDS knowledge measure for adolescents focusing on misconceptions. *Journal of HIV/AIDS Prevention in Children and Youth*, *8*(1), 35–57.

Bruine de Bruin, W., Fischhoff, B., Halpern-Felsher, B., *et al*. (2000). Expressing epistemic uncertainty: it's a fifty-fifty chance. *Organizational Behavior and Human Decision Processes*, *81*, 115–31.

Bruine de Bruin, W., Güvenç, Ü., Fischhoff, B., Armstrong, C.M., and Caruso, D. (2009). Communicating about xenotransplanation: models and scenarios. *Risk Analysis*, *29*, 1105–15.

Bruine de Bruin, W., Parker, A., and Fischhoff, B. (2007b). Individual differences in adult decision-making competence (A-DMC). *Journal of Personality and Social Psychology*, *92*, 938–56.

Budescu, D.F. and Wallsten, T.S. (1995). Processing linguistic probabilities: general principles and empirical evidence. In J.R. Busemeyer, R. Hastie, and D.L. Medin (eds.) *Decision Making from the Perspective of Cognitive Psychology*, pp. 275–316. New York: Academic Press.

Byram, S., Fischhoff, B., Embrey, M., *et al*. (2001). Mental models of women with breast implants regarding local complications. *Behavioral Medicine*, *27*, 4–14.

Campbell, P. (2011). Understanding the receivers and the receptions of science's uncertain messages. *Philosophical Transactions of the Royal Society*, *369*, 4891–912.

Christensen-Szalanski, J. and Bushyhead, J. (1993). Physicians' misunderstanding of medical findings. *Medical Decision Making*, *3*, 169–75.

Dawes, R.M., Faust, D., and Meehl, P. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–74.

DeKay, M.L., Small, M.J., Fischbeck, P.S., *et al*. (2002). Risk-based decision analysis in support of precautionary policies. *Journal of Risk Research*, *5*, 391–417.

Downs, J.S., Bruine de Bruin, W., and Fischhoff, B. (2008). Patients' vaccination comprehension and decisions, *Vaccine*, *26*, 1595–607.

Downs, J.S., Bruine de Bruin, W., Murray, P.J., *et al*. (2004b). When 'it only takes once' fails: perceived infertility predicts condom use and STI acquisition. *Journal of Pediatric and Adolescent Gynecology*, *17*, 224.

Downs, J.S., Murray, P.J., Bruine de Bruin, W., *et al*. (2004a). An interactive video program to reduce adolescent females' STD risk: a randomized controlled trial. *Social Science and Medicine*, *59*, 1561–72.

Eggers, S.L. and Fischhoff, B. (2004). Setting policies for consumer communications: a behavioral decision research approach. *Journal of Public Policy and Marketing*, *23*, 14–27.

Erev, I. and Cohen, B.L. (1990). Verbal versus numerical probabilities: efficiency, biases and the preference paradox. *Organizational Behavior and Human Decision Processes*, *45*, 1–18.

Ericsson, K.A. and Simon, H.A. (1993). *Verbal Reports as Data*. Cambridge, MA: MIT Press.

Fagerlin, A. and Peters, E. (2011). Quantitative information. In B. Fischhoff, N.T. Brewer, and J.S. Downs (eds.) *Communicating Risks and Benefits: An Evidence-Based User's Guide*, pp. 53–64. Washington, DC: US Food and Drug Administration.

Farris, C. and Fischhoff, B. (2012). A decision science informed approach to sexual risk and non-consent. *Clinical and Translational Science, 5*, 482–5.

Feather, N. (1982). *Expectancy, Incentive and Action*. Hillsdale, NJ: Erlbaum.

Finucane, M.L. and Gullion, C.M. (2010). Developing a tool for assessing the decision-making competence of older adults. *Psychology & Aging*, *25*, 271–88.

Fischhoff, B. (1992). Giving advice: decision theory perspectives on sexual assault. *American Psychologist*, *47*, 577–88.

Fischhoff, B. (1994). What forecasts (seem to) mean. *International Journal of Forecasting*, *10*, 387–403.

Fischhoff, B. (1995). Risk perception and communication unplugged: twenty years of process. *Risk Analysis*, *15*, 137–45.

Fischhoff, B. (1996). The real world: what good is it? *Organizational Behavior and Human Decision Processes*, *65*, 232–48.

Fischhoff, B. (2005a). Cognitive processes in stated preference methods. In K.G. Mäler and J. Vincent (eds.) *Handbook of Environmental Economics*, pp. 937–68. Amsterdam: Elsevier.

Fischhoff, B. (2005b). Decision research strategies. *Health Psychology*, *21*, S9–16.

Fischhoff, B. (2008). Assessing adolescent decision-making competence. *Developmental Review, 28*, 12–28.

Fischhoff, B. (2011). Communicating the risks of terrorism (and anything else). *American Psychologist, 66*, 520–31.

Fischhoff, B., Brewer, N.T., and Downs, J.S. (eds.) (2011). *Communicating Risks and Benefits: An Evidence-Based User's Guide*. Washington, DC: US Food and Drug Administration.

Fischhoff, B., Bruine de Bruin, W., Guvenc, U., *et al*. (2006). Analyzing disaster risks and plans: an avian flu example. *Journal of Risk and Uncertainty*, *33*, 133–51.

Fischhoff, B., Downs, J., and Bruine de Bruin, W. (1998). Adolescent vulnerability: a framework for behavioral interventions. *Applied and Preventive Psychology*, *7*, 77–94.

Fischhoff, B. and Kadvany, J. (2001). *Risk: A Very Short Introduction*. Oxford: Oxford University Press.

Fischhoff, B. and MacGregor, D. (1983). Judged lethality: how much people seem to know depends upon how they are asked. *Risk Analysis*, *3*, 229–36.

Fischhoff, B., Parker, A., Bruine de Bruin, W., *et al*. (2000). Teen expectations for significant life events. *Public Opinion Quarterly*, *64*, 189–205.

Fischhoff, B., Slovic, P., and Lichtenstein, S., (1977). Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 552–64.

Fischhoff, B., Slovic, P., Lichtenstein, S., *et al*. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, *8*,

127–52.

Fischhoff, B., Watson, S., and Hope, C. (1984). Defining risk. *Policy Sciences*, *17*, 123–39.

Florig, K. and Fischhoff, B. (2007). Individuals' decisions affecting radiation exposure after a nuclear event. *Health Physics*, *92*, 475–83.

Florig, H.K., Morgan, M.G., Morgan, K.M., *et al*. (2001). A deliberative method for ranking risks. *Risk Analysis*, *21*, 913–22.

Fortney, J. (1988). Contraception: a life long perspective. In *Dying for Love*, pp. 33–8. Washington, DC: National Council for International Health.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Funtowicz, S.O., and Ravetz, J. (1990). *Uncertainty and Quality in Science for Policy*. London: Kluwer.

Gilovich, T., Griffin, D., and Kahneman, D. (eds.) (2003). *Judgment Under Uncertainty II: Extensions and Applications*. New York: Cambridge University Press.

Griffin, D., Gonzalez, R., and Varey, C. (2003). The heuristics and biases approach to judgment under uncertainty. In A. Tesser and N. Schwarz (eds.) *Blackwell Handbook of Social Psychology*, pp. 207–35. Boston, MA: Blackwell.

HM Treasury (2005). *Managing Risks to the Public*. London: HM Treasury.

Howard, R.A. (1989). Knowledge maps. *Management Science*, *35*, 903–22.

Kahneman, D., Slovic, P., and Tversky, A. (eds.) (1982). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Koriat, A. (1993). How do we know that we know? *Psychological Review*, *100*, 609–39.

Krimsky, S. and Golding, D. (1992). *Theories of Risk*. New York: Praeger.

Lerner, J.S. and Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*, 146–59.

Lichtenstein, S. and Fischhoff, B. (1980). Training for calibration. *Organizational Behavior and Human Performance*, *26*, 149–71.

Lichtenstein, S., Fischhoff, B., and Phillips, L.D. (1982). Calibration of probabilities. In D. Kahneman, P. Slovic, and A. Tversky (eds.) *Judgment Under Uncertainty: Heuristics and Biases*, pp. 306–39. New York: Cambridge University Press.

Lichtenstein, S., Slovic, P., Fischhoff, B., *et al*. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, *4,* 551–78.

Linville, P.W., Fischer, G.W., and Fischhoff, B. (1993). AIDS risk perceptions and decision biases. In J.B. Pryor and G.D. Reeder (eds.) *The Social Psychology of HIV Infection*, pp. 5–38.

Hillsdale, NJ: Erlbaum.

Löfstedt, R., Fischhoff, B., and Fischhoff, I. (2002). Precautionary principles: general definitions and specific applications to genetically modified organisms (GMOs). *Journal of Policy Analysis and Management*, *21*, 381–407.

Lowrance, W.W. (1976). *Of Acceptable Risk: Science and the Determination of Safety*. Los Altos, CA: William Kaufman.

McIntyre, S. and West, P. (1992). What does the phrase 'safer sex' mean to you? Understanding among Glaswegian 18 year olds in 1990. *AIDS*, *7*, 121–6.

Merton, R.F. (1987). The focussed interview and focus groups. *Public Opinion Quarterly*, *51*, 550–66.

Merz, J., Fischhoff, B., Mazur, D.J., *et al*. (1993). Decision-analytic approach to developing standards of disclosure for medical informed consent. *Journal of Toxics and Liability*, *15*, 191–215.

Meyer, D., Leventhal, H., and Gutmann, M. (1985). Common-sense models of illness: the example of hypertension. *Health Psychology*, *4*, 115–35.

Morgan, M.G., Fischhoff, B., Bostrom, A., *et al*. (1992). Communicating risk to the public. *Environmental Science and Technology*, *26*, 2048–56.

Morgan, M.G., Fischhoff, B., Bostrom, A., *et al*. (2001). *Risk Communication: The Mental Models Approach*. New York: Cambridge University Press.

Murphy, A.H., Lichtenstein, S., Fischhoff, B., *et al*. (1980). Misinterpretations of precipitation probability forecasts. *Bulletin of the American Meteorological Society*, *61*, 695–701.

National Research Council (1989) *Improving Risk Communication*. Washington, DC: National Academy Press.

National Research Council (1996). *Understanding Risk: Informing Decisions in a Democratic Society*. Washington, DC: National Academy Press.

National Research Council (2006). *Scientific Review of the Proposed Risk Assessment Bulletin from the Office of Management and Budget*. Washington, DC: National Academy Press.

Nickerson, R.A. (1999). How we know—and sometimes misjudge—what others know: imputing our own knowledge to others. *Psychological Bulletin*, *125*, 737–59.

O'Hagan, A., Buck, C.E. Daneshkhah, A., *et al*. (2006). *Uncertain Judgements: Eliciting Expert Probabilities*. Chichester: Wiley.

Peters, E. and McCaul, K.D. (eds.) (2005). Basic and applied decision making in cancer. *Health Psychology*, *24*(4), S3.

Politi, M.C., Han, P.K.J., and Col. N. (2007). Communicating the uncertainty of harms and benefits of medical procedures. *Medical Decision Making*, *27*, 681–95.

Poulton, E.C. (1989). *Bias in Quantifying Judgment*. Hillsdale, NJ: Lawrence Erlbaum.

Quadrel, M.J., Fischhoff, B., and Davis, W. (1993). Adolescent (in)vulnerability. *American Psychologist*, *48*, 102–16.

Reimer, B., and Van Nevel, J.P. (eds.) (1999). Cancer risk communication. *Journal of the National Cancer Institute Monographs*, *19*, 1–185.

Reyna, V. and Farley, F. (2006). Risk and rationality in adolescent decision making: implications for theory, practice, and public policy. *Psychology in the Public Interest*, *7*(1), 1–44.

Riley, D.M., Fischhoff, B., Small, M., *et al*. (2001). Evaluating the effectiveness of risk-reduction strategies for consumer chemical products. *Risk Analysis*, *21*, 357–69.

Ritov, I. and Baron, J. (1990). Status quo and omission bias. Reluctance to vaccinate. *Journal of Behavioral Decision Making*, *3*, 263–77.

Schwartz, L.M. and Woloshin, S. (2011). Communicating uncertainties about prescription drugs to the public: a national randomized trial. *Archives of Internal Medicine*, *171*, 1463–8.

Schwarz, N. (1999). Self reports. *American Psychologist*, *54*, 93–105.

Shaklee, H. and Fischhoff, B. (1990). The psychology of contraceptive surprises: judging the cumulative risk of contraceptive failure. *Journal of Applied Psychology*, *20*, 385–403.

Slovic, P. (2001). *Perception of Risk*. London: Earthspan.

Slovic, P. and Fischhoff, B. (1983). Targeting risk. *Risk Analysis*, *2*, 231–8.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1978). Accident probabilities and seat-belt usage: a psychological perspective. *Accident Analysis and Prevention*, *10*, 281–5.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1979). Rating the risks. *Environment*, *21*(4), 14–20, 36–9.

Slovic, P., Lichtenstein, S., and Fischhoff, B. (1984). Modeling the societal impact of fatal accidents. *Management Science*, *30*, 464–74.

Slovic, P., Peters, E., Finucane, M.L., *et al*. (2005). Affect, risk and decision making. *Health Psychology*, *24*, S35–40.

Sox, H.C., Blatt, M.A., Higgins, M.C., *et al*. (2007). *Medical Decision Making*. Philadelphia, PA: American College of Physicians.

Starr, C. (1969). Societal benefit versus technological risk. *Science*, *165*, 1232–8.

Sterman, J. and Sweeney, J. (2002). Cloudy skies: assessing public understanding of climate change. *System Dynamics Review*, *18*, 207–40.

Thaler, R. (1991). *Quasi-Rational Economics*. New York: Russell Sage Foundation.

Thaler, R. and Sunstein, C. (2009). *Nudge: Improving Decisions about Health, Wealth and Happiness.* New Haven, CT: Yale University Press.

USEPA (1993). *A Guidebook to Comparing Risks and Setting Environmental Priorities*. Washington, DC: USEPA.

Viscusi, K. (1992). *Smoking: Making the Risky Decision*. New York: Oxford University Press.

Vlek, C. and Stallen, P.J. (1981). Judging risks and benefits in the small and in the large. *Organizational Behavior and Human Performance*, *28*, 235–71.

Von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioral Research*. New York: Cambridge University Press.

Wagenaar, W. and Sagaria, S.D. (1975). Misperception of exponential growth. *Perception & Psychophysics*, *18*, 416–22.

Willis, H.H., DeKay, M.L., Fischhoff, B., *et al*. (2005). Aggregate and disaggregate analyses of ecological risk perceptions. *Risk Analysis*, *25*, 405–28.

Wogalter, M. (2006). *The Handbook of Warnings*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Woloshin, S., Schwartz, L.M., Byram, S., *et al*. (1998). Scales for assessing perceptions of event probability: a validation study. *Medical Decision Making*, *14*, 490–503.