IMPROVING COMPARATIVE RISK ANALYSIS

JAMES K. HAMMITT[*]

Comparative Risk Analysis (CRA) was introduced by the U.S. Environmental Protection Agency (EPA) as a method for allocating agency resources across environmental issues. The purpose of the initial EPA staff CRA, *Unfinished Business,*[1] and of the follow-up study by the EPA Science Advisory Board, *Reducing Risk,*[2] was to compare the relative magnitudes of threat posed by environmental factors under EPA jurisdiction with the agency staff, budget, and attention allocated to each factor. The initial report found that EPA resources were not allocated primarily toward what its staff viewed as the largest risks, but that EPA resource allocations followed public perceptions of the most significant risks, as determined from public-opinion polling. The CRA study found that hazardous air pollutants, indoor radon, and global-scale issues like climate change and stratospheric-ozone depletion ranked lower in terms of public perception and agency priority than the agency staff CRA suggested they should. Following efforts by Administrators William Ruckelshaus and Lee Thomas to orient EPA activities toward risk reduction, Administrator William Reilly used CRA as a central element in developing a process of risk-based priority setting within the agency.[3]

Building on these path-breaking studies, the ten EPA regional offices, together with many states, municipalities, Indian nations, and other jurisdictions, have conducted their own CRAs.[4] The format of these exercises has varied substantially, both from the initial EPA efforts and among each other. In particular, the state and local studies

---

[*] Associate Professor, Center for Risk Analysis and Department of Health Policy and Management, Harvard School of Public Health.

1. U.S. ENVIRONMENTAL PROTECTION AGENCY, OFFICE OF POLICY ANALYSIS, UNFINISHED BUSINESS: A COMPARATIVE ASSESSMENT OF ENVIRONMENTAL PROBLEMS 8 (1987).

2. U.S. ENVIRONMENTAL PROTECTION AGENCY, SCIENCE ADVISORY BOARD, REDUCING RISK: SETTING PRIORITIES AND STRATEGIES FOR ENVIRONMENTAL PROTECTION 12 (1990).

3. *See* F. Henry Habicht III, *EPA's Vision for Setting National Environmental Priorities, in* WORST THINGS FIRST? THE DEBATE OVER RISK-BASED NATIONAL ENVIRONMENTAL POLICIES, 33, 37-41 (Adam M. Finkel & Dominic Golding eds., 1994); Charles W. Kent & Frederick W. Allen, *An Overview of Risk-Based Priority Setting at EPA, in* WORST THINGS FIRST?, *supra*, at 47, 50-53.

4. *See* Richard A. Minard, *CRA and the States: History, Politics, and Results, in* COMPARING ENVIRONMENTAL RISKS: TOOLS FOR SETTING GOVERNMENT PRIORITIES 23, 33-34 (1996).

81

have involved much greater public participation than the EPA studies. Whereas the federal EPA studies were conducted by agency staff and other experts, many of the more recent studies have included significant numbers of outsiders, including members of the general public, as well as representatives of industries, environmental advocacy groups, and other "stakeholder" groups. This shift in participation appears to reflect the recognition that ranking environmental risks requires a substantial input of information on human preferences and values, topics of which risk assessment experts have little specific expertise.[5] It may also reflect a difference in objective, with the subsequent studies directed more toward developing a political consensus than were the initial EPA studies.

As reflected in the diversity of approaches undertaken, there is not yet a consensus on how a CRA should be conducted, which is not surprising given the recent development of the field and its complexity. In "Ranking Risks: Some Key Choices", Clarence J. Davies provides an overview of some of the key design issues confronting CRAs.[6] To provide useful guidance for conducting a CRA, this article examines three critical design issues: (1) what kind of items are ranked; (2) who participates in the analysis and what information is used in determining the rankings; and (3) the extent to which rankings should be constrained by economic and other decision-making principles. The following section discusses how the choices made with respect to each of these design issues should depend on the goals of the CRA. Whether risks or opportunities for risk reduction should be the items ranked is discussed in Section II. Section III considers what expertise participants in a CRA should have and what information should be available to them. Section IV introduces some principles of decision making and examines their implications for CRA.

---

5. *See* U.S. ENVIRONMENTAL PROTECTION AGENCY, *supra* notes 1-2.

6. J. CLARENCE DAVIES, *Ranking Risks: Some Key Choices, in* COMPARING ENVIRONMENTAL RISKS: TOOLS FOR SETTING GOVERNMENT PRIORITIES 9 (J. Clarence Davies ed., 1996).

I. Goals of a CRA Should Affect Its Design

As the Cheshire Cat pointed out to Alice, it does not matter what path you take if you do not care where you go. Similarly, the best design for a CRA will depend on its goals.

CRAs may be (and have been) conducted in the furtherance of several goals, including allocating government or social resources to environmental threats, building a political consensus around a vision of environmental protection, the developing of information, and educating of government officials, stakeholders, public-opinion leaders, risk analysts, and citizens. The scope of the analysis can vary both in the domain of risks included and in the temporal perspective. A micro-level analysis could compare risks of alternative disinfection technologies for drinking water, or could guide the allocation of EPA resources among hazardous air pollutants or toward the selection of compounds for cancer bioassay through the National Toxicology Program. The early EPA reports are macro-scale, comparing risks from global climate change, hazardous waste, indoor and outdoor air pollution, drinking and surface waters, and other sources. The EPA analyses were restricted to risks over which the EPA has at least some statutory authority; an even broader CRA might consider a full range of risks to human health, safety, and the environment. Similarly, the temporal perspective of the analysis might range from setting an agency's budget or program activities for the next year to sustainable development of the biosphere.

Appropriate design choices will depend on the goals of a CRA. Figure 1 provides a stylized mapping from goal to choices on three primary design criteria. Three possible goals of a CRA are considered: efficient resource allocation, development of a political consensus, and broader education of policy makers and citizens. Within each of the three design characteristics, two choices are offered: participants may be technical experts or general citizens, the items ranked may be risks or actions (opportunities for risk reduction), and rankings may or may not be constrained by appeal to exogenous decision-making principles. While the illustrated goals and design options are only caricatures, Figure 1 highlights some of the possible dependence of study design on goals.

Figure 1: Illustrative dependence of design choices on goal of CRA

	Design Characteristic					
	Items Ranked		Principles		Participation	
Goal of Analysis	*Risks*	*Actions*	*Experts*	*Public*	*Constrain*	*Open*
Resource Allocation		x	x		x	
Political consensus	x	x		x		x
Education	x	x		x	x	x

In this article, "efficient resource allocation" means the maximization of risk reduction within resource limits. The goal of efficient resource allocation is likely to be best approached using a fairly technical CRA, conducted by experts rather than by members of the general public, ranking available actions or opportunities for risk reduction rather than risks, and relying on decision-analytic principles to facilitate and ensure the consistency of the ranking. To some extent, agreement on the goal of efficient risk reduction resolves issues of preferences, in that risk reduction must be defined for the exercise to proceed. For example, risk could be defined in terms of lifetime mortality risks from selected environmental causes or quality-adjusted life years (QALYs) lost. To the extent that other factors are to be included in measuring risk, inputs can be obtained through existing research on the structure of public preferences; for example, the substantial economic and medical literature on valuation of health risk and environmental quality.[7] The federal EPA analyses are examples of this approach, although more formal, quantitative assessments could also be undertaken.[8]

If the primary goal is to achieve political consensus or at least a shared vision of the direction in which environmental, health, and safety regulation should proceed within a political jurisdiction, it is essential that elected representatives, stakeholders, and even non-specialist members of the public be directly involved in generating the rankings. Choices about the other design characteristics are

---

7. *See* A. MYRICK FREEMAN III, THE MEASUREMENT OF ENVIRONMENTAL AND RESOURCE VALUES: THEORY AND METHODS (1993); W. Kip Viscusi, *The Value of Risks to Life and Health,* 31 J. OF ECON. LITERATURE 1912 (1993); George Torrance, *Measurement of Health State Utilities for Economic Appraisal: A Review,* 5 J. OF HEALTH ECON. 1 (1986); MARTHE R. GOLD ET AL., COST-EFFECTIVENESS IN HEALTH AND MEDICINE (1996).

8. *See* U.S. ENVIRONMENTAL PROTECTION AGENCY, *supra* notes 1-2.

more flexible; it may be useful to rank either risks or actions (or both), and principles of consistency may be given more or less attention as desired. To the extent that a workable political agreement which permits expeditious regulatory decisions is more important than achieving some theoretical notion of optimality, the details of what is agreed to become less important. Note that it may be possible to achieve agreement on a ranking of actions without achieving agreement on a ranking of risks, or vice versa. Participants may be able to agree on priorities among actions because they value different attributes of the actions, such as risk reduction, job creation, symbolism, and others; indeed, a great deal of legislative action involves agreement (or at least compromise) on means without agreement on desired ends. Alternatively, agreement on ranking risks might generate some progress in resolving debates over actions by resolving or at least removing some fundamental value issues from debate, allowing the focus to turn to the question of how best to mitigate those risks.

If the goal is simply one of educating the general public, stakeholder groups, and/or elected representatives, again the choices of what items are ranked and what principles imposed are less important. The arguments for ranking actions rather than risks and for imposing principled constraints on preferences deserve more attention when the goal is education than when it is consensus building, because a goal of education is to enhance the sophistication with which policy makers evaluate future risk decisions. Many of the state and local CRAs seem to have given higher priority to the consensus-building and educational goals, and have consequently included more public participation than the early national EPA studies.[9]

II. RANKING RISKS OR RISK-REDUCTION OPPORTUNITIES?

The original federal CRAs and most of the subsequent studies have attempted to rank the risks resulting from various pollutants and other sources that exist given current levels of control, often designated "residual risks." While ranking risks is appealing at first glance, it is neither necessary nor sufficient for allocating resources to minimize exposure to risk. Moreover, there is no apparent basis for categorizing risks, and the categorization selected can affect the ranking. In contrast, ranking risk-reduction opportunities overcomes

---

9. *See* Minard, *supra* note 4, at 33-35.

these problems, at least in principle. In practice, however, ranking actions requires much more information than does ranking risks, and therefore ranking risks may be a useful substitute.

In order to maximize risk reduction given limited resources, risks should be ranked by cost-effectiveness of available actions. Figure 2 illustrates the magnitude of risk reduction achievable, as a function of cost, for three hypothetical risks. Under the current resource allocation, A poses the largest risk, followed by B then C. Assume a small quantity of additional resources can be devoted to reducing one or more of these risks. As illustrated, the maximum risk reduction will be achieved by spending these resources on risk C, because the marginal cost-effectiveness of spending is larger than for either of the other risks (the risk-cost curve is steeper for C than for the other risks; marginal cost-effectiveness is the absolute value of the slope of this curve). Marginal cost-effectiveness is larger for risk A than for B, so the correct ordering for minimizing risk would be C, A, B.

A second problem with ranking risks, rather than actions, is that the ranking may depend on how risks are classified; there is no rationale for selecting the appropriate typology. Risks can be classified by the physical or chemical agent responsible (e.g., pesticides, ionizing radiation), by the human or natural activity generating the risk (e.g., production of electricity, transportation), by the exposure pathway (e.g., air, drinking water), or by other criteria. Moreover, the appropriate degree of aggregation is indeterminate.[10] For example, air pollution might rank relatively high, but if the class were disaggregated into tropospheric ozone, particulates, carbon monoxide, and other hazardous pollutants, some of these would rank high and others low. Particulates might be further disaggregated between large and small particulates (greater and smaller than 2.5 microns in diameter, for example); small particulates would likely rank as a more significant risk than large particulates.[11]

---

10. *See* John D. Graham et al., *Refining the CRA Framework, in* COMPARING ENVIRONMENTAL RISKS, *supra* note 6, at 93, 98.

11. *See* C. Arden Pope III et al., *Health Effects of Particulate Air Pollution: Time for Reassessment?*, 103 ENVTL. HEALTH PERSP. 472 (May 1995); Mary Amdur et al., *Animal Toxicology, in* PARTICLES IN OUR AIR: CONCENTRATIONS AND HEALTH EFFECTS 85 (Richard Wilson & John Spengler eds., 1996).

Figure 2: Comparing risks and risk-reduction opportunities
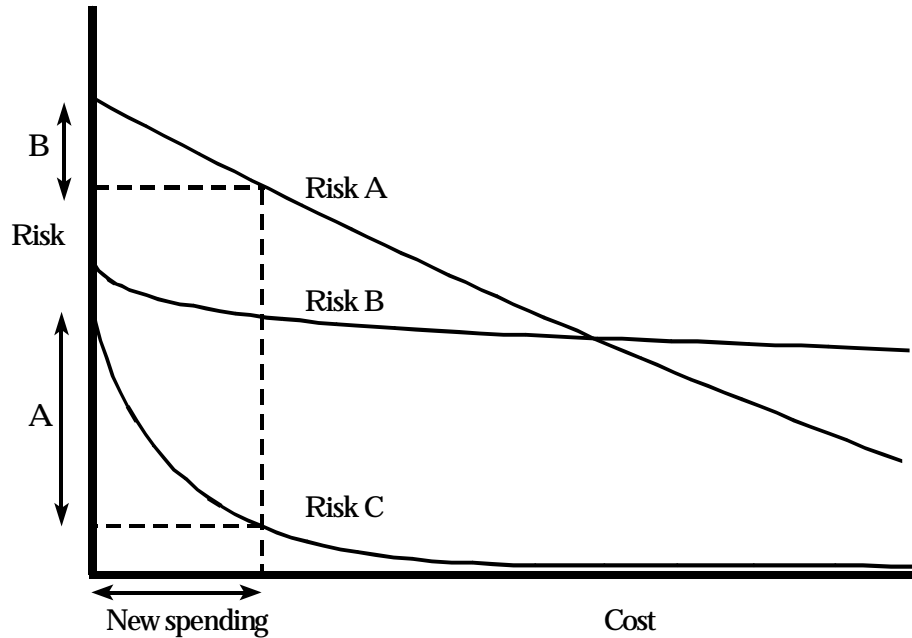
In contrast, ranking risk-reduction opportunities provides several advantages. Problems of aggregation can be resolved by preliminary analysis of the marginal or incremental cost-effectiveness of opportunities and aggregating those with similar cost-effectiveness. The aggregation can be subsequently refined if desirable. In principal, all possible opportunities can be analyzed individually and arranged in order of decreasing marginal cost-effectiveness to produce risk-abatement curves like those illustrated in Figure 2. As an example, Professor Tammy O. Tengs, et al. report cost-effectiveness of more than 500 life-year-saving interventions involving control of toxins, injury prevention, and medicine.[12]   In addition, analysis of risk-reduction options allows one to incorporate effects of actions on both the target risk and any countervailing or offsetting risks exacerbated by the action.[13]

---

12. *See* Tammy O. Tengs et al., *Five-Hundred Life-Saving Interventions and Their Cost-Effectiveness,* 15 RISK ANALYSIS 369 (1995).

13. *See* JOHN D. GRAHAM & JONATHAN BAERT WIENER, *Confronting Risk Tradeoffs, in* RISK VERSUS RISK: TRADEOFFS IN PROTECTING HEALTH AND THE ENVIRONMENT 1, 10-17 (John D. Graham & Jonathan Baert Wiener eds., 1995).

The drawback to ranking risk-reduction opportunities is that it requires substantially more information than does ranking risks. In addition to evaluating population exposure to risk factors and exposure/risk relationships, it is necessary to consider the efficacy of risk reduction technologies, countervailing risks, and costs of control. Ranking risks can serve as a useful proxy and analytic short-cut to ranking options if risk magnitude and marginal cost-effectiveness are well correlated across risks (i.e., if the curves that start high in a plot like Figure 2 are also the ones that fall most steeply). This condition will tend to hold when exposures of populations to risks can be reduced at similar costs, for example.

The choice between ranking risks and risk-reduction opportunities is an example of the general problem of choosing an evaluation endpoint. In many environmental-policy contexts, it is impossible to reliably estimate the effect of different risk factors or policies on endpoints that are of direct human concern; consequently, policies are evaluated in terms of some proxy factor that is assumed to be related to the endpoints of interest and can be projected with greater confidence. In regulating possible chemical carcinogens, for example, the relevant endpoint might be the distribution over the population of the incremental probability of developing different forms of cancer at various ages; in practice, a lifetime risk of any type of cancer is used instead. In evaluating the risk of stratospheric-ozone depletion due to chlorofluorocarbons (CFCs), relevant endpoints include the increases in human skin cancer, ocular damage, and effects on wildlife from enhanced ultraviolet radiation. Beginning with the 1974 discovery that CFCs might deplete ozone, total column ozone loss (an idealized global average measure) was used as the evaluation endpoint, on the assumption that greater ozone loss would allow more ultraviolet radiation to reach the Earth's surface. Discovery of the Antarctic ozone hole made clear (a) that existing scientific models did not adequately explain global ozone patterns; and (b) that seasonal and geographic variation in ozone depletion was much greater than had been believed.[14] The implication was that total column ozone was both less accurately projected and less relevant to evaluating ultraviolet exposure to humans and wildlife than had been thought. Following this discovery, analyses of ozone depletion have typically substituted the concentrations of chlorine and bromine in

---

14. *See* J.C., Farman et al., *Large Losses of Total Ozone in Antarctica Reveal Seasonal ClO$_x$/NO$_x$ Interaction,* 315 NATURE 207 (1985).

the stratosphere (the agents responsible for ozone depletion) as the evaluation endpoint.[15]

III. PARTICIPATION AND INFORMATION

Health and environmental risks typically involve many attributes. In addition to the probabilities of death, morbidity, and damage to ecosystems, people may care about the extent of pain and suffering, the distribution of risk among people and whether it corresponds to the distribution of benefits from the risk-generating activity, perceived familiarity and controllability of the risk, and other factors.[16] Moreover, individual preferences about risks and their association with these factors may differ depending on available information, experience, political outlook, and other factors. The multidimensional nature of risk and heterogeneity in preferences raise issues about whose preferences should be used in the CRA, and what information should risk-rankers have available.

Developing a social ordering of risks is necessarily a matter of compromise. In his famous "impossibility theorem," Kenneth J. Arrow showed that there is no social ordering of goods that satisfies four seemingly reasonable criteria:

- Sensitivity to individual preferences. If one individual shifts his preference between A and B so that he ranks B superior to A, the social ranking should not be changed to rank A above B.

- Non-dictatorship. The social ordering is not determined by a single individual.

- Insensitivity to individual preferences. If the social ordering ranks A above B above C, removing B from the set does not produce an ordering with C ranked above A.

---

15. *See* James K. Hammitt & Kimberly M. Thompson, *Protecting the Ozone Layer, in* THE GREENING OF INDUSTRY: A RISK MANAGEMENT APPROACH 43, 84-85 (John D. Graham & Jennifer K. Hartwell eds., 1997).

16. *See* Paul Slovic et al., *Facts Versus Fears: Understanding Perceived Risk, in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 463 (Daniel Kahneman et al. eds., 1982); Paul Slovic*, Perception of Risk,* 236 SCIENCE 280 (1987); Baruch Fischhoff, *Risk: A Guide to Controversy, in* IMPROVING RISK COMMUNICATION 211 (National Research Council et al. eds., 1989).

- Universal domain. The social ordering can be determined for all possible combinations of individual rankings.[17]

For social allocations of benefits, it is conventional to distinguish efficiency from distribution. An allocation is (Pareto) efficient if there is no reallocation such that no one in society prefers his lot before the reallocation than after. Typically, a very large number of allocations are efficient; in the problem of dividing a birthday cake among partygoers, any division that leaves no cake on the serving plate is efficient, regardless of whether the whole cake goes to one gluttonous guest or is evenly divided among all of them. Efficiency is not, of course, the only goal, and people are often willing to give up some efficiency for greater equity or other values.[18]

Knowledge and access to information about risks are likely to vary widely within a population. In performing a risk-ranking exercise, it seems reasonable to propose that all relevant and reliable information about the risks should be provided to participants, subject only to concerns about over-burdening them with information of secondary relevance. Judgments about relevance and reliability are likely to differ, however. In the health-utility literature, preferences over various disease and disability conditions are often characterized using numerical values on a scale between zero (dead) and one (perfect health). Conditions of ill health are often viewed as less bad by people who have those conditions than they are viewed by unaffected people, perhaps because healthy people underestimate the extent to which they could adapt to disease or disability.[19] This suggests that a particular disease may rank differently depending on whether the participants are those who suffer from the disease, those who do not suffer from it, or those who do not suffer from it but who have close contact with (or are otherwise informed by) those who do suffer from it.

Reliability of information can also be controversial. In particular, habits of framing scientific information can lead to different conclusions about reliability. Consider a potential risk for which there is limited evidence as to whether or not the risk even exists, such as the carcinogenic effect of low-level exposure to certain chemicals. As-

---

17. KENNETH J. ARROW, SOCIAL CHOICE AND INDIVIDUAL VALUES (1963).

18. *See* ARTHUR M. OKUN, EQUALITY AND EFFICIENCY: THE BIG TRADEOFF (1975); Peter A. Ubel et al., *Cost-Effectiveness Analysis in a Setting of Budget Constraints,* 334 NEW ENG. J. MED. 1174 (1996).

19. MARTHE R. GOLD ET AL., *supra* note 7, at 100.

sume we wish to classify the chemical as carcinogenic if and only if the probability that it will induce human cancer exceeds one per million (for some specified exposure scenario). Applying classical (frequentist) statistics to the problem, the conclusion will depend on what hypothesis is selected as the null or default.

This risk is not measurable using any practical experiment, and so it is correct to say that there is no direct evidence of (non-negligible) risk, but also no direct evidence of negligible risk. In a classical statistical hypothesis-testing analysis, the conclusion about whether the risk exists or not may depend on how the problem is framed: if the chemical is to be assumed non-hazardous unless there is evidence of risk, it will be found safe; if the null hypothesis is that the chemical is hazardous unless evidence shows the risk to be smaller than some criterion, it will be found hazardous. Conclusions about the magnitude of the risk are necessarily subject to extrapolations that cannot be adequately tested.

IV. PRINCIPLED CONSTRAINTS ON PREFERENCES

One way to define a consistent pattern of choices is in terms of its adherence to a set of axioms of choice. Axioms of choice are statements about preferences that capture primitive and intuitively compelling aspects of what we mean by consistency; choices violating one or more accepted axioms would not be regarded as consistent.

If a specified set of risks (or of risk-reduction actions) can be ranked, then the ranking must satisfy two axioms: asymmetry and negative transitivity. Asymmetry means that if A is ranked before B, B cannot be ranked before A.[20] At least in the abstract, the desirability of complying with these axioms seems to be compelling, and the fact that a ranking must satisfy both of them is intuitive. What is perhaps less intuitive is that if rankings among a pre-specified set of risks satisfy both axioms, then a complete ranking can be obtained. Such a ranking is called a "preference relation." Moreover, the ranking can be represented by an ordinal utility function: a function that assigns a numerical value to each risk such that one risk is ranked higher than another if and only if its utility exceeds the utility of the other risk.[21]

---

20. Negative transitivity means that if A is not ranked before B and B is not ranked before C, then A is not ranked before C. Equivalently, if A is ranked before B, C must be ranked before B or after A (or both).

21. *See* DAVID M. KREPS, NOTES ON THE THEORY OF CHOICE (1988).

Because this utility function is ordinal (meaning only the order of values counts, not their differences or ratios), any convenient scale can be chosen. One obvious scale is the rank itself (a utility function can be defined so that a smaller rank is more preferred). Other scales could use monetary units, QALYs, or other metrics (additional axioms beyond asymmetry and negative transitivity would be required for the differences in numbers of dollars or QALYs to be meaningful).

The existence of a complete ranking implies that tradeoffs are made among the attributes of each of the risks, and that preferences over the attributes are revealed (at least in part) by choices. For example, if the loss of 100 million acres of wetland were ranked worse than the extinction of one species of migratory songbird, which was ranked worse than the loss of 1 million acres of wetland, one could infer that the appropriate tradeoff between acres of wetland loss and extinction of songbirds is at a ratio somewhere between 1 million and 100 million to one.

Tradeoffs among quantitative and qualitative risk attributes have been examined by various scholars.[22] Professors Maureen L. Cropper and Uma Subramanian used nationwide survey data to examine preferences among lives saved from various public health and environmental risks, characteristics of the risks, and characteristics of the government programs. In hypothetical choices among programs saving various numbers of lives, they found that choices depended on characteristics of the risks and of the programs. Respondents were more likely to select a program saving fewer lives if the risk from which the lives were saved was more serious, less controllable, and if the respondent was exposed to the risk. Respondents were also more likely to choose a program saving fewer lives if the program was viewed as likely to be effective, if it was likely to be an appropriate intervention for government to undertake, if the funding was perceived as fair, and if the program would yield benefits relatively

---

22. *See* MAUREEN L. CROPPER & UMA SUBRAMANIAN, PUBLIC CHOICES BETWEEN LIFESAVING PROGRAMS: HOW IMPORTANT ARE LIVES SAVED? (The World Bank, Policy Research Working Paper No. 1497, August 1995); Timothy L. McDaniels, *Comparing Expressed and Revealed Preferences for Risk Reduction: Different Hazards and Question Frames,* 8 RISK ANALYSIS 593 (1988); Timothy L. McDaniels et al., *Risk Perception and the Value of Safety,* 12 RISK ANALYSIS 495 (1992); Ian Savage, *An Empirical Investigation into the Effect of Psychological Perceptions on the Willingness-to-Pay to Reduce Risk,* 6 J. OF RISK AND UNCERTAINTY 75 (1993); W. Kip Viscusi et al., *Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-Offs for Chronic Bronchitis,* 21 J. OF ENVTL. ECON. & MGMT. 32 (1991).

quickly. The effects of differences in risk and program attributes were quantitatively significant: in some cases, the median respondent would require twice as many lives to be saved by one program (e.g., reducing radon exposure in homes) to prefer it to another program (either a workplace smoking ban or a pesticide ban on fruit). Moreover, as many as 30 to 40% of respondents were estimated to prefer one program to another saving 20 to 100 times as many lives (e.g., smoking education versus control of industrial air-pollution; colon-cancer screening versus control of drinking-water pollution).

A number of potential violations of choice axioms have been identified that may be relevant to a CRA. In some cases, the violation appears to be unacceptable: there is no compelling set of axioms that would justify the violation. In others, the normative significance is less clear: the axioms may fail to capture important aspects of a decision, and therefore their violation is justified. Four potential violations are described below.

A. *Framing Effects*

Professors Amos Tversky and Daniel Kahneman have identified a number of apparent violations of choice axioms that they explain as systematic results of reliance on heuristic decision rules. One of the most prominent and troubling of these violations is the effect of framing: human choice may depend critically on the way in which a question is posed, even when the alternative questions are logically equivalent.[23] In one well-known example, Tversky and Kahneman posed the following problem:

- Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. . . .

- [Lives Saved:] If program A is adopted, 200 people will be saved. . . . If program B is adopted, there is 1/3 probability that 600 people will be saved, and 2/3 probability that no people will be saved. . . .

---

23. *See* Amos Tversky & Daniel Kahneman, *Judgments Under Uncertainty: Heuristics and Biases,* 185 SCIENCE 1124 (1974); Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk,* 47 ECONOMETRICA 291 (1979).

- [Lives Lost:] If program C is adopted 400 people will die. . . .
  If program D is adopted there is 1/3 probability that nobody
  will die, and 2/3 probability that 600 people will die.[24]

Responses to the different question frames (by different samples
of people) are clearly inconsistent. In the first question frame, 72%
of respondents choose A over B, saving 200 people for sure rather
than gambling on saving all of them with one-third probability.   In
the second framing, almost the same fraction—78%—prefer D to C,
gambling on saving all 600 rather than letting 400 die for sure.[25] The
existence of framing effects poses an important challenge to CRA,
since it is important to ensure that rankings are not sensitive to logi-
cally equivalent ways of presenting the risks.

B. *Sensitivity to Scope*

Tradeoffs between attributes, whether implicit or explicit, should
generally depend on the magnitudes of the changes in attribute lev-
els. The amount of a valued attribute that one should be willing to
give up for an improvement in another attribute should depend on
the magnitude of the improvement. This assumption has received ex-
tensive scrutiny in the context of giving up money (willingness to pay
or WTP) for improvements in health and environmental quality. In
particular, the validity of a prominent method of eliciting preferences
for diverse goods—Contingent Valuation—has been questioned be-
cause the resulting estimates of WTP do not seem to be sufficiently
sensitive to the magnitude of the improvement in environmental
quality.

Contingent Valuation (CV) attempts to elicit preferences by
asking survey respondents either to state the maximum amount of
money they would pay for a certain specified change in environ-
mental conditions or other goods, or to choose between two options
differing in cost and environmental quality. In some cases, CV-based
estimates of WTP for environmental improvements do not appear to
be sufficiently sensitive to the magnitude of the improvement. In one
example, respondents were asked their WTP (in the form of higher
prices) for wire-net covers to protect migratory waterfowl from being

---

24. Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, 59
J. OF BUS. 251, 260 (1986).

25. *Id.*

killed by landing in waste-oil holding ponds in the Central Flyway (a region including parts of Texas, Oklahoma, and New Mexico). Different subsets of respondents were told that the covers would prevent different numbers of birds from dying each year: 2,000, 20,000, and 200,000 (the numbers were supplemented by descriptions as "much less than 1%," "less than 1%," and "about 2%" of the 8 million waterfowl migrating along this route annually). The estimated WTP for the three levels of environmental improvement were $59, $59, and $71, respectively.[26] The small differences in WTP compared with the 100-fold variation in number of birds saved leads some commentators to suggest that CV did not accurately measure tradeoffs between wealth and environmental quality in this instance.[27]

It is not clear exactly how WTP should relate to the number of birds protected in this example. A naive assumption might be that WTP should be proportional to the number of birds saved, yielding a 100-fold difference in WTP between the 2,000 and 200,000 bird scenarios. The usual assumption of diminishing marginal rates of substitution would suggest that WTP should increase less than in proportion to the number of birds saved. Considering the number of living birds as the relevant attribute of the situation, rather than the number of birds saved, suggests that WTP should be increase more than proportionately to the number of birds saved, since the scenarios then represent increases in population to 8,000,000 from 7,998,000, 7,980,000, and 7,800,000, respectively.[28] Alternatively, the attribute of greatest concern may not be the number of birds affected but rather the fact that free-living birds are being killed in waste-oil ponds; in that case, perhaps WTP should not be sensitive to the number of birds affected.

For human mortality risks, economic theory provides a stronger justification for the appropriate tradeoff between risk reduction and WTP. If the marginal value of wealth is greater to an individual if he is alive than if he is dead, WTP for a small reduction in mortality risk (within the next year or other time period) is positive and should be nearly proportional to the magnitude of the reduction in mortality probability. For otherwise similar individuals with different levels of

---

26. *See* William H. Desvousges et al., *Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability, in* CONTINGENT VALUATION: A CRITICAL ASSESSMENT 91 (Jerry A. Hausman ed., 1993).

27. *Id; see also* Peter A. Diamond & Jerry A. Hausman, *Contingent Valuation: Is Some Number Better than No Number?*, 8 J. OF ECON. PERSP. 45, 51-52 (1993).

28. *See* Peter A. Diamond, *Testing the Internal Consistency of Contingent Valuation Surveys* 30 J. OF ENVTL. ECON. AND MGMT. 337 (1996).

wealth or different levels of mortality risk within some time period, the one with the larger wealth or the larger period-mortality risk should have larger WTP for the same risk reduction. WTP may fall substantially below proportionality to the risk reduction only when the reduction is so large that the payment causes substantial loss of income.[29]

C. *Risk Aversion and Equity*

In some cases, seemingly desirable attributes are in direct conflict. Such is the case with a desire for equitable risk distributions within a population and risk aversion over the number of deaths from the risk.

Consider two distributions of mortality risk, summarized in Figure 3. In case A, each of the 1,000 residents of a community face probabilistically independent 1/1,000 mortality risks from a specified cause. In case B, two residents face a 1/2 mortality risk from a specified cause; the other 998 residents face no risk. Assuming that the benefits of whatever activity produces the risk are widespread, Case A appears to be far more equitable; case B imposes extremely large risks on two unfortunate individuals.

In both cases, the expected number of deaths from the risk is one. The worst-case outcomes are very different: in case B, the worst outcome is two deaths; in case A, all 1,000 people could die from the risk (although the probability is vanishingly small—10^{-3000}). The probability that no one dies from the risk is substantially larger in case A than in case B (0.37 and 0.25, respectively). The probability that more than two people die from the risk is zero in case B but almost 0.08 in case A; however, the probability that more than four people die in case A is less than 0.003. Taking into account only the probabilities of numbers of deaths, not the interpersonal distribution of risks, case B appears to be more attractive; technically, it can be shown that risk aversion over the number of deaths (preferring N deaths for certain to a gamble with N expected deaths) implies a preference for B over A. More generally, Professor Ralph Keeney has shown that a preference for equitable distribution of risks and

---

29. *See* Michael Jones-Lee, *The Value of Changes in the Probability of Death or Injury,* 82 J. OF POL. ECON. 835 (1974); Milton C. Weinstein et al., *The Economic Value of Changing Mortality Probabilities: A Decision-Theoretic Approach,* 94 Q. J. OF ECON. 373 (1980); John W. Pratt & Richard J. Zeckhauser, *Willingness to Pay and the Distribution of Risk and Wealth,* 4 J. OF POL. ECON. 747 (1996).

risk-aversion over the number of deaths are fundamentally incompatible: a preference for equity implies a risk-seeking preference over the number of deaths.[30]

Figure 3: Alternative Distributions of Mortality Risk

	Probability	
Deaths	**A: 1,000 people face independent 1/1,000 risk**	**B: 2 people face independent ½ risk, 998 face no risk**
0	0.37	0.25
1	0.37	0.50
2	0.18	0.25
3	0.06	0
4	0.015	0
≥ 5	< 0.003	0

D. *Uncertainty Aversion and Learning from Experience*

Studies of the factors affecting tolerance or acceptability of risks have often found that uncertainty about the risk, sometimes in the form of unfamiliarity about technology, makes a risk less acceptable.[31] In decision theory, uncertainty aversion leads to the Ellsberg Paradox: people prefer a lottery with known odds over a similar lottery with unknown odds.[32] Consider two choices of gambles involving which color ball will be drawn from an urn containing 300 balls, of which 100 are known to be red and 200 are known to be blue or yellow (but the division between blue and yellow balls is unknown). Given the choice of winning $1,000 if a red ball or a blue ball is drawn, many respondents prefer to bet on red. Given the choice of winning $1,000 if either (a) a red or yellow ball is drawn, or if (b) a blue or yellow ball is drawn, the same respondents often prefer to bet on blue or yellow, however.

---

30. *See* Ralph Keeney, *Utility Functions for Equity and Public Risk*, 26 MGMT. SCI. 345 (1980).

31. *See* Paul Slovic, *Perception of Risk,* 236 SCIENCE 280 (1987).

32. *See* Daniel Ellsberg, *Risk, Ambiguity, and the Savage Axioms,* 75 Q. J. OF ECON. 643 (1961).

As shown in Figure 4, such preferences appear to reflect inconsistent estimates of the mixture of blue and yellow balls in the urn. If one thinks there are more blue than yellow balls, the probability P of drawing a blue ball is greater than 1/3, so one should bet on the second choice in each gamble (blue; blue or yellow); alternatively, if one thinks there are more yellow than blue balls, he should bet on the first choice in each gamble. If, as seems reasonable, he has no idea whether there are more blue than yellow balls, he should not have a strong preference on which side of the bets he takes.

Figure 4: The Ellsberg Paradox

|	**Ball Drawn**			
Probability	Red	Blue	Yellow	
	1/3	P	2/3 - P	
Bet on:				Prob(win)
Red	$1,000	$0	$0	1/3
Blue	$0	$1,000	$0	P
Bet on:				Prob(win)
Red or Yellow	$1,000	$0	$1,000	1 - P
Blue or Yellow	$0	$1,000	$1,000	2/3

In many cases, uncertainty about a risk can be a reason to choose that risk over a more certain one, if exposure to the uncertain risk yields information that is useful for subsequent decisions. Consider the choice between an innovative, unfamiliar technology and a standard, reliable technology for remediating Superfund sites. Assume both technologies cost the same to apply. The innovative technology will work with probability P; if so, it will completely remove the health and environmental risks associated with the site (risk reduction = 1); if it fails, it will yield no risk reduction (risk reduction = 0). The standard technology reduces the risk by an amount V, somewhere between 0 and 1.

If the goal of the program is to maximize the cumulative risk reduction across sites, the appropriate choice between the innovative and standard technologies will depend on the number of future sites at which these technologies will be applicable. If there is only one site at which the decision is to be made, the expected risk reduction is

maximized by choosing the innovative technology if and only if P exceeds V. If the success of the innovative technology at the first site is a perfect indicator of whether it will work at future sites, it is worth choosing the innovative technology at the first site to obtain this information. If the technology works, it can then be used with confidence at all future sites where it is applicable; if it fails, the standard technology should be used at all future sites.

The value of information obtained by trying the innovative technology can be quite large. If the probability of success is one-half and there is only one additional site where the new technology may be used, it should be selected for the first site if V < 2/3. If there are nine additional sites at which the new technology is applicable, it should be tried unless the standard technology provides almost complete remediation (V > 10/11 ≈ 0.91). As shown by Figure 5, if there are more than 10 or so sites at which this decision must be made, it is better to choose the innovative technology at the first site unless the probability that it will succeed is small, or the remediation provided by the standard technology is nearly complete. The larger the number of sites involved, the more valuable it is to learn that the innovative technology works, and therefore the bigger the risk of failure one should accept on the first try.[33]

Figure 5: Choosing Between Innovative and Standard Technologies
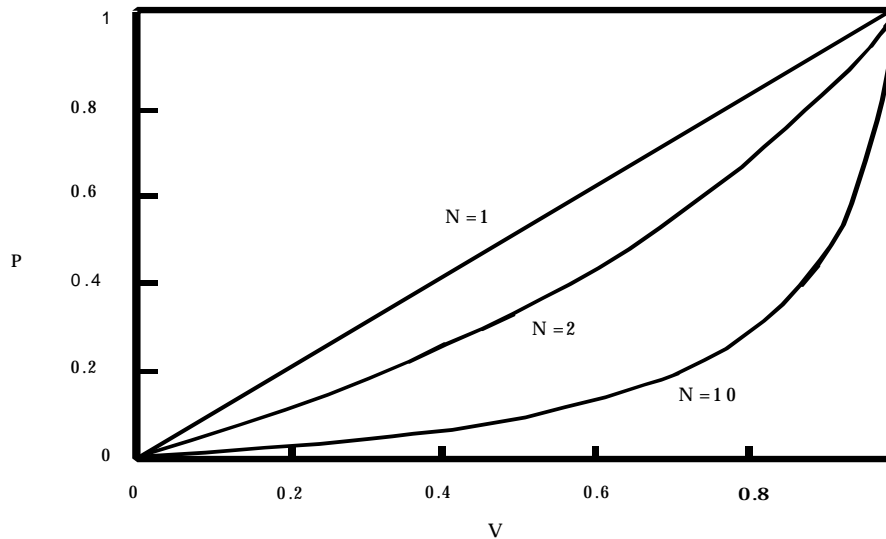
33. For N sites total, the innovative technology should be selected if $P > [1 + N (1 - V)/V]^{-1}$

V. SUMMARY

Comparative Risk Analysis encompasses a range of study designs addressing a broad set of goals. Several key design issues have been considered, including whether risks or risk-reduction actions are the items that are ranked; the mix of participants between technical risk assessors and representatives of the general public; and the extent to which rankings should be informed by decision-analytic principles. The best design for a CRA will depend on the goals it is intended to address, but in many cases it will be useful to at least consider the implications of decision-analytic principles. Violations of some principles, as in the sensitivity of choices to problem framing, appear to be unjustifiable. Care should be taken in conducting CRAs to test for and alleviate these violations. Other violations, such as apparent inadequate sensitivity to scope, may be justified, but the findings of a CRA will be strengthened if such violations are identified and justified through reasoned argument rather than overlooked in haste or ignorance.