

From Comparative Risk to Decision Analysis: Ranking Solutions to Multiple-Value Environmental Problems*

John Kadvany**

*Science begins and ends with problems.*¹
Karl Popper

Priority-Setting As a Problem

Environmentalists, regulators, industry personnel and concerned citizens have a basic interest in how to set or negotiate environmental priorities given limited and possibly changing resources. Problems include selecting wetlands for costly and uncertain restoration projects, cleaning up hazardous waste sites to various land use standards or earmarking funds for pollution prevention. They are partly technical or scientific in that solutions require some knowledge of the physical, biological or social conditions associated with various risks. They are also normative in that decision makers must address what should be done with available or additional resources — or what should be delayed or not done at all.

A challenge of environmental priority-setting is not only of ranking risks, but also *solutions to risk problems*. Even a single type of problem, such as remediating hazardous waste sites, has multiple values associated with it such as human health risk, ecological risk or groundwater contamination (considered as a natural resource); a variety of socioeconomic impacts such as property transfer or environmental

* For encouragement and comments, I thank Christina Chociolko, Baruch Fischhoff, Karen Jenni, Paul Slovic and two reviewers. Bill Soo Hoo, Jeff Wong, Kimi Klein, Lee Merkhofer, Karen Wong and Antje Kann made essential contributions to the case study described below.

** Dr. Kadvany is a Senior Manager at Applied Decision Analysis, Inc. in Menlo Park, California. He received his Ph.D. from the Group in Logic and the Methodology of Science at the University of California, Berkeley.

¹ Quoted in Imre Lakatos, *Proofs and Refutations: The Logic of Mathematical Discovery* 5 (1976).

equity; and costs such as that of restoring wetlands. Priority-setting entails trade offs among these competing values when resources are inadequate to do everything; resource consumption demands prudence, or additional resources require negotiation.

"Values" refer here to why we are more supportive of certain decision outcomes than others or what we care about, e.g., human health and safety, ecological quality, the quality and extent of natural resources, aesthetic values such as visibility and economic impacts. Yet, values take on significance only within a decision-making context; *without the need to choose or prioritize actions, values have little meaning*. Conversely, any approach to value modeling that does not help you decide what to do is seriously deficient.

Ranking risks (problems), is obviously part of ranking solutions: You cannot assess the adequacy of alternative solutions without understanding the problem they are intended to address.

Comparative Risk

Consider the approach of the U.S. Environmental Protection Agency (EPA) to ranking environmental problems. This arose after its 1987 study, **Unfinished Business**, concluded that environmental policy was often guided too much by public and political perceptions that poorly correlated with risk levels and that social resources were not being allocated to reduce health and environmental risk most effectively. This then led to 1990 EPA Science Advisory Board (SAB) efforts to, e.g.:²

² See U.S. Environmental Protection Agency (EPA), **Unfinished Business: A Comparative Assessment of Environmental Problems** (1987), **Reducing Risk: Setting Priorities and Strategies for Environmental Protection** (and appendices, 1990); and **A Guidebook to Comparing Risks and Setting Environmental Priorities** (1993). Several of ten recommendations of **Reducing Risk**, at 6, imply that more is needed than ranking problems, bear repeating:

1. EPA should target its environmental protection efforts on the basis of opportunities for the greatest risk reduction. ...
4. EPA should reflect risk-based priorities in its strategic planning process.
5. EPA should reflect risk-based priorities in its budget process. ...
7. EPA should emphasize pollution prevention as the preferred option for reducing risk.
8. EPA should increase its efforts to integrate environmental considerations into broader aspects of public policy in as fundamental a manner as are economic concerns. ...
10. EPA should develop improved analytical methods to value natural

1. base rankings on the best available knowledge while reflecting uncertainties;
2. address environmental (i.e. ecological and natural resource) as well as health impacts;
3. consider the social and economic impacts associated with different risks;
4. make explicit the connections between technical knowledge and value-laden policy judgments;
5. provide means for addressing views of multiple stakeholders; and
6. consider solutions such as pollution prevention as well as more traditional approaches.

Since at least then, the comparison of multiple value outcomes has been part of EPA's agenda, as was the intention to somehow dovetail solutions with risk problems. While neither the original SAB reports nor subsequent EPA materials on comparative risk created a complete approach to risk and solution rankings, the pieces of the puzzle were articulated and their combined importance for policy recognized. In particular, the straightforward idea of generating policy recommendations by looking at solutions and their costs, in tandem with problems, was there. EPA's comparative risk paradigm has resulted in over a dozen completed ranking projects in Vermont, California, Washington, Colorado and other states.³ Over twenty further projects, sometimes involving dozens of participants, are underway or planned as of mid-1995. These projects represent one of EPA's most sustained efforts in public participation and policy-setting. The largest, such as the California project completed in 1994, cost nearly \$1M. Yet, while these comparative risk projects have allowed EPA to better incorporate values and tradeoffs in setting policy, the original goals set out in 1990 have been only partially achieved, and some have apparently been forgotten.

Let's briefly evaluate some of the strengths and weaknesses of its comparative risk projects. Identifying and characterizing uncertainties may be the most successful aspect of most of them. For example, although using a somewhat ambiguous "high," "medium," "low" resources and to account for long-term environmental effects in its economic analyses.

³ See **Comparative Risk Bulletin** (Northeast Center for Comparative Risk) (July/August 1995), at 7-9.

classification, the California project provided "level of confidence" judgments with detailed descriptions for about two dozen health hazard sources and structured likelihood judgments for dozens of ecological exposure routes.

Competing sources of value, typically those of health, environment and social welfare are also identified and individually ranked in several projects. Unfortunately, the meaning of these multiple rankings and their comparative analysis is relatively informal and has not proven compelling. For example, the integrated rankings of the Vermont study do not appear to be definable independently of participants' judgments at ranking sessions, and the whole dimension of value is discussed largely as a personal and subjective matter that apparently could not be articulated with adequate definiteness.⁴

As in the original SAB report, the term "importance" is used as a proxy in several reports for more careful conceptions of value. Also, in many cases, while the right issues such as the extent of health risk, relative harm and populations of concern (e.g., children versus adults)⁵ have been raised, the ambiguities of the term "importance" make it almost impossible to disentangle technical and value judgments, or to determine just what value judgments were intended.

There has also been little development of measurement scales used to rank risks. For example, the Vermont "quality of life" scale and the California "social welfare" scale combine attributes as disparate as aesthetics, economic well-being, peace of mind, fairness, future generations, recreation and sense of community. Combining all these dimensions without some further structure makes it difficult to know what such scales are intended to represent, and the ultimate policy influence that such rankings could have is diminished.

Stakeholders have had many opportunities to participate, and the projects represent one of EPA's most successful efforts in public participation to date. Yet, the public and other stakeholders typically have had few opportunities to define the environmental problems considered, include new value categories or consider novel solutions to

⁴ See *Environment 1991: Risks to Vermont and Vermonters* (State of Vermont) (1991) at 15.

⁵ For examples of "importance" see *Environment 1991 supra* at 15 and *Reducing Risk, supra* note 2, at 17. See also *infra* note 26 and associated text.

risk problems. In the Vermont report, for example, it was concluded that the problem of "people and their consumption of resources pose the greatest risk to Vermont and Vermonters." However, this was not officially part of the analysis: The greatest risk was "considered" but "not ranked!"⁶ As "overconsumption" was relegated to ambiguous status in the Vermont report, the 1994 California report treats economics and environmental justice on tracks parallel to the apparently more "official" rankings of health, ecology and social welfare.⁷

A stumbling block in comparative risk appears to be at the step of creating an integrated ranking of risks, or in the explicit treatment of multiple and competing values. Some projects stop short at an integrated ranking (e.g., California), but, when integrated rankings are created, the logic tends to be ad hoc and largely intractable. Thus, it would be difficult to obtain a clear answer to the question, "Well, what do you mean by saying you 'considered human health risk and ecological risk to be equally important?'"⁸ Such approaches are a consequence of EPA's offering little cogent advice in comparative risk documents to explain what *integrated rankings should represent and how they should be constructed*.

Moreover, as Baruch Fischhoff asked recently, "What can one do with a risks ranking?"⁹ From the perspective of action, risk ranking produces a "worst-first" list and imply that one should first address top-ranked risks. Still, if one of your top-ranked risks is global warming, as in the Vermont report, should Vermont pursue that to the exclusion of all others until it's fixed? By leaving out EPA's original intention to consider solutions in tandem with problems, a risk ranking

⁶ See *Environment 1991 supra* note 4, at 3, 36.

⁷ Importantly, all the reports are neither one-sided in defining what "really" is a problem nor in sharply segregating problems from solutions. The 1990 Washington State report, for example, proactively defines priorities for pollution prevention and other actions, and has been used to help develop state legislation. See *Toward 2010: An Environmental Action Agenda* (State of Washington) (1990).

⁸ See *Reducing Risk supra* note 2 at 6: "EPA should attach as much importance to reducing ecological risk as it does to reducing human health risk."

⁹ Baruch Fischhoff, *Ranking Risks*, 6 *Risk* 191, 201(1995). Note also the advice of the National Research Council, *Improving Risk Communication*, 174-175 (1989) that risk comparisons can help us understand likelihoods, gauge the relative importance of different causes of the same hazard and help compare available options. Comparative risk has largely excluded the third goal.

may also fail to identify causes that can lead to effective risk reduction. Also, because "problem rankings" serve no traditional scientific goal, it's unclear who is served by the process.

In summary, comparative risk generally does much better with treating uncertainty and codifying scientific information with which EPA and its affiliates are familiar, as opposed to incorporating values — probably because of the central role of uncertainty in risk assessment. Comparative risk is well-motivated by the introduction of the issues and structure defined by the SAB in 1990, as reflected in projects' multiple rankings and the use of stakeholder input. However, once critical problems are raised, i.e., comparing multiple and competing values and how a ranking is to be used, the method fails to provide substantive guidance.

Project and EPA documentation incorrectly suggest that this is the best that could be done with current knowledge about integrating science and values. Thus, comparative risk has been largely self-limited to ranking problems and generally fails to bring technical and value judgments together in policy recommendations. Along the way, EPA's original motivation to include pollution-preventing alternatives, economics of environmental policy and merging of problems and solutions has not been sustained, and EPA's approach is not solving the problems it set for itself. When so many individuals have obviously contributed so much time and worked so hard to develop these reports, it may seem churlish to press methodological issues.

Still, the question has to be raised: To what extent may methodological failings make difficult work less efficient and contain less clear purpose than possible — and ultimately be less effective than it could be in influencing environmental policy? Perhaps consider another approach would fare better.

Ranking Problem Solutions

Benefit represents not only the size of a problem addressed, but also the difference between pursuing specific remedial action and not doing so. The idea behind ranking solutions is to compare benefits expected from, e.g., a remediation project, to that achieved by not doing it. Assume for the purpose of argument that we know how to define

"project benefit" meaningfully, to incorporate whatever multiple value considerations are deemed relevant, such as reducing health risk, environmental risk and adverse socioeconomic impacts. Then the ranking is defined by estimating the total benefit of the project (e.g., site remediation) and dividing this value by project costs; that is the benefit-to-cost measure. Note that this doesn't integrate risks, as in comparative risk, but rather *reduced risks*, or benefits, from solutions.

A key confusion in comparative risk surrounds the role and meaning for "integration." While it's difficult to meaningfully integrate a set of diverse risks without reference to risk reduction, if we value the *elimination* of risk, such values, as shown above, can be integrated.

Table 1¹⁰ ranks, according to benefit-to-cost ratio, with benefits expressed in dollars,¹¹ site studies or cleanup actions administered by a large U.S. regional program. The largest benefit-cost project is at the top, the smallest at the bottom. An illustrative, though overly crude use of such a ranking, is to simply "draw a budget line" where cumulative costs are just less than budget, assuming perhaps artificially that your only choices to conduct or delay these activities: Given resources, you would fund the projects above the point at which cumulative costs do not exceed a given budget and delay, or not fund, projects below the line. Depending on context, you might also argue that the proposed budget excludes worthwhile projects and that it should be increased.

¹⁰ The benefits and costs in Table 1 are expressed in \$M. The benefit for conducting site studies, for which risk is left unchanged, is based on a "value-of-information" calculation representing the expected cost or risk reduction attributable to the study. The relevant value weights are derivable from the health, ecological and groundwater categories; no additional value weight is used. For value-of-information, see Robert Clemen, *Making Hard Decisions* 339 ff (1991); for a policy discussion, see Jeffrey Harris, *Environmental Policy Making: Act Now or Wait for More Information?* in National Research Council, *Valuing Health Risks, Costs, and Benefits for Environmental Decision Making* 107 (1990).

¹¹ The data shown are for a pilot project conducted during 1994. Project data, the model structure and value weights are yet to be reviewed and made public. Since a motivation for the approach presented is that the whole ranking is entirely open and revisable, data errors or inconsistencies in value judgments are not here an issue, though obviously they should be dealt with.

For a summary of a similar system for evaluating environmental restoration projects for the Department of Energy's nuclear weapons complex, see National Research Council, *Ranking Hazardous Waste Sites for Remedial Action* ch. 6 (1994) and Office of Technology Assessment, *Complex Cleanup: The Environmental Legacy of Nuclear Weapons Production* 60 ff (1991).

Table 1
Example of Prioritized List of Projects
(Site Studies and Cleanups)

Project #	Project Description	Health & Safety Benefit	Ecological & GW Benefit	Value of Uncertainty Reduction	Socio-Economic Benefit	Total Benefit	Project Cost	B/C Ratio	Cumulative Cost	Cumulative Benefit
1	Cleanup	1.06	1.6		0.1	2.8	0.3	11.0	0.3	2.8
2	Site Study			1.9	8.6	10.5	1.2	8.4	1.5	13.3
3	Cleanup	0.004	0.2		1.0	1.1	0.2	5.1	1.7	14.4
4	Cleanup	0.05	2.1		0.1	2.3	0.5	4.3	2.2	16.7
5	Site Study			0.8	0.3	1.1	0.3	4.3	2.5	17.8
6	Cleanup				1.0	1.0	0.3	3.7	2.8	18.8
7	Cleanup	0.02			3.1	3.1	1.0	3.1	3.8	21.9
8	Site Study			1.0	2.1	3.1	1.1	3.0	4.8	25.0
9	Cleanup	0.002	0.7		1.6	2.4	0.6	2.6	5.4	26.7
10	Site Study			0.2	1.0	1.2	0.5	2.3	6.0	27.9
11	Site Study			1.0	0.4	1.5	0.9	1.6	6.9	29.3
12	Site Study			0.003	7.7	7.7	5.1	1.5	12.0	37.0
13	Site Study			79.3		79.3	53.0	1.5	65.0	116.4
14	Cleanup				1.0	1.0	0.7	1.3	65.7	117.3
15	Site Study			0.5	0.2	0.7	0.6	1.2	66.3	118.0
16	Cleanup	0.03	1.3			1.3	1.4	0.9	67.7	119.3
17	Cleanup	0.25			0.4	0.6	1.0	0.7	68.7	120.0
18	Site Study	0.004			2.0	2.0	3.9	0.5	72.6	122.0
19	Site Study			1.3	1.1	2.4	4.6	0.5	77.2	124.4
20	Cleanup	0.04			1.9	1.9	4.0	0.5	81.2	126.3
21	Cleanup	0.004	1.3		2.1	3.4	7.9	0.4	89.1	129.7
22	Site Study			1.5		1.5	4.9	0.3	94.0	131.2
23	Site Study				1.1	1.1	4.2	0.3	98.2	132.3
24	Cleanup	0.07			2.0	2.1	82.8	0.02	181.0	134.3
25	Site Study			1.2		0.2	8.9	0.02	189.9	134.5
26	Site Study			0.1		0.1	4.0	0.02	193.9	134.6

Figure 1
Benefit-Cost Curve for 26 Site Remediation Projects

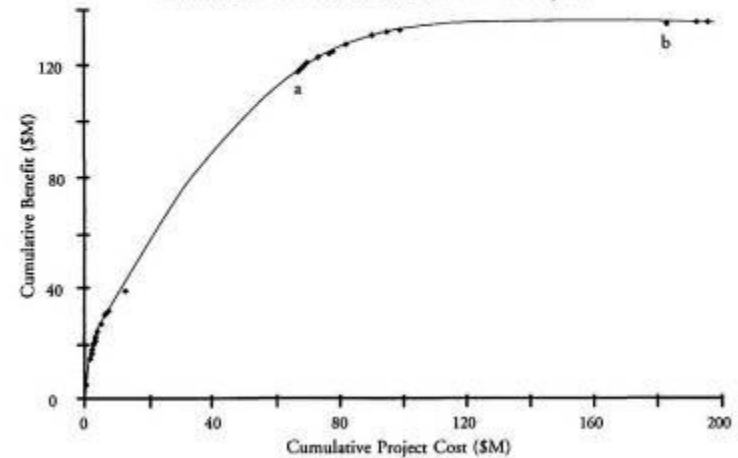


Figure 1 shows projects taken in priority order from left to right, each contributing its incremental benefit on the Y-axis and its incremental cost on the X-axis. The curve shows the distribution, over projects, of potential total benefits versus cumulative total costs. All else being equal, this analysis implies that you maximize total benefit with a constrained budget by conducting projects from left-to-right (or "top-down" in Table 1) in priority order. Projects at the lower-left in Figure 1 are high-benefit-to-cost and might be funded first from the budget at issue, all else being equal; the high-benefit-and-high-cost project at (a) may be valuable, but its scale may cause difficulties with respect to more typical smaller endeavors. At the upper right, (b) is a relatively low-benefit-and high-cost project which you might want to delay or not funding at all.

One can also simply invent new options and see how they rank. Thus, you might come up with lower-cost alternatives to (b), such as site studies or cleanups based on the most realistic (as opposed to conservative) future land use assumptions, as long discussed by EPA and others. In this way, the ranking is used to interrogate risks and the

projects addressing them, but no conclusions follow mechanically. Basically, you try to identify projects or project options that, *ceteris paribus*, are most worthwhile, the factors responsible — defensibly and openly. It is important that ranking is relative and comparative, not absolute. You might well want to argue for the importance of funding all projects, depending on resources (e.g., historically underfunded or historically inflated), the nature of the projects (e.g., wetland restorations vs. hazardous waste sites) and the overall purpose of priority-setting (e.g., reallocating resources from clean ups to prevention). Priority-setting is not synonymous with triage.¹²

Because project costs, by no means certain, are typically least controversial in project rankings, the key to this approach is defining project benefit with multiple value outcomes. Given that you have a meaningful measure of multiple-value benefit, the remainder of the analysis for ranking solutions is relatively straightforward, as shown here. Next, we examine briefly how a multiple-values benefit model is developed using the methods of multiattribute utility theory, the foundation of the approach.

Value Modeling Using Multiattribute Utility Theory

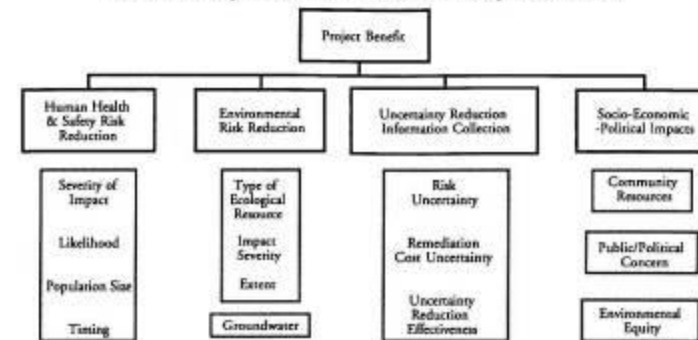
Multiattribute utility theory (MUA) is an axiomatized mathematical framework for analyzing choices involving multiple competing outcomes.¹³ Briefly, multiattribute utility theory is formalized using four fundamental ideas.

¹² It is tempting to use a B/C ratio such as 1 as a cutoff point based on the argument that, if the B/C is less than 1, the marginal return is less than your marginal investment. That may be reasonable if the benefit measure is "well-calibrated," but that assumption is difficult to justify in practice given the approximate nature of inputs. Generally it is best to assume that rankings are relatively consistent but that benefits may need to be benchmarked against standards external to the analysis. On program expenditures and health risk reduction, see, e.g., John D. Graham & James W. Vaupel, *Value of a Life: What Difference Does It Make?* 1 *Risk Anal.* 89 (1981).

¹³ See Detlof von Winterfeldt & Ward Edwards, *Decision Analysis and Behavioral Research* (1986); Ralph L. Keeney, *Value-Focused Thinking* (1992). For a discussion of MUA and ranking solutions, see M. Granger Morgan, *Quantitative Risk Ranking: More Promise Than the Critics Suggest*, 133 in *Worst Things First? — The Debate over Risk-Based National Environmental Priorities* (Adam M. Finkel & Dominic Golding, eds. 1994).

1. A Value or Objectives Hierarchy

Figure 2
Value Hierarchy Associated with Prioritized Projects of Table 1



A value (or objectives) hierarchy organizes matters of interest, such as minimizing ecological risk, minimizing health risk, minimizing various socioeconomic impacts, costs, etc. A *value* is, e.g., "health risk," measured appropriately; its associated *objective* is "minimize health risk." An example of such a value hierarchy, defining the project rankings listed in Table 1, is shown in Figure 2.¹⁴ The value categories include human health and safety risk, ecological risk, groundwater contamination, community resource property transfer, public and political concern, and environmental equity.

2. Measurement Scales and Utility Functions

Individual measurement scales are used to evaluate how well a specific alternative (e.g., a removal action or characterization study) is expected to perform on a single objective. Such scales might use formal risk assessments, professional judgments of experts, consensus judgments of a citizen panel or some mixture of these.¹⁵ In the

¹⁴ Inputs used to calculate benefits in distinct categories are shown in light-border boxes. Some benefit categories use just one input (e.g., groundwater) while others (health, ecological, uncertainty reduction) use multiple inputs. Note that without an explicit calculation for uncertainty reduction, site studies might receive no benefit.

¹⁵ On public participation and MUA, see Ralph L. Keeney, Detlof von Winterfeldt & Thomas Eppel, *Eliciting Public Values for Complex Policy Decisions*, 36 *Management Sci.* 1011 (1990).

example given, formal health risk assessments were used when available, while ecological impact, impacts on community resources (basically property transfer) and other inputs were judged by managers familiar with sites and the projects being conducted using detailed scales appropriate for the sites and projects evaluated.¹⁶ Measurement scales are also used to define "single-attribute utility functions" representing the desirability of an alternative with respect to a *single* objective; these functions provide the first step in translating technical inputs into value-laden scores that will in turn be combined in a single measure of multiattribute utility or benefit.

3. Value or Policy Weights

Value or policy weights are associated with each distinct value, representing the relative tradeoffs one is willing to make, in the decision context, between different outcomes. A set of weights associated with the hierarchy, expressed in program dollars, is shown below in Table 2.¹⁷

¹⁶ An example of a measurement scale is provided below in Table 3.

¹⁷ The value weights below were neither internally nor externally reviewed and would likely be revised after such a process. Comments in Table 2 indicate how the weights and model structure might change depending on stakeholder views.

Definitions were provided for "poor," "marginal," "median," "good" and "protected" habitats by participating ecotoxicologists. For example, the definition of "median" is:

Habitat which while suboptimal, may be expected to support native species, or commercially important species, in addition to species which are proposed, recommended, or under review for State or Federal listing as rare, threatened or endangered, or which are State species of special concern. Areas utilized or potentially utilized by wildlife, including but not limited to sanctuaries, preserves, easements, wildlife rehabilitation center, or game management areas.

For environmental equity, the instructions read:

Will members of a disadvantaged ethnic, racial, or socioeconomic group potentially be affected by the proposed activity in terms of environmental equity? 'Environmental inequity' refers to the disproportionate distribution of health or environmental risks to disadvantaged groups, or to intentional or unintentional biases in how risks to such groups have been addressed. Environmental equity depends significantly both on local demographics, the history of government and other official actions, and comparable patterns elsewhere.

This definition was intended largely as a place-holder for a better characterization; the modularity of the MUA model makes such submodeling tasks manageable. Land use decisions, typically associated with many equity disputes, were also not within the scope of the program sponsoring development of the model.

Table 2
Value Weights Associated with Rankings and Value Hierarchy in Table 1 & Fig. 1

Value Category	Relative Value (in Program \$)	Summary Meaning and Comments
Human Health & Safety	\$5M	Relative value of averting one "most serious" (e.g., premature cancer) statistical health impact. This category is not intended to apply to any imminent, non-statistical impacts. The same value weight is used for worker and public impacts, but worker impacts attributable to remediation were not included in estimates.
Ecological	\$2.6M	Relative value of avoiding destruction of ten acres of "median" ecological habitat.
Groundwater	\$1.8M	Relative value of avoiding contamination of groundwater resource associated with site. Does not reflect either extent or quality of resource.
Community Resources	\$950K	Relative value of making available an average site intended to be developed for commercial or residential use. Deliberately not scaled by acreage or other dimensions to represent program indifference regarding the importance of different development projects.
Environmental Equity	\$950K	Relative value of completing a project judged to contribute to environmental equity. This input should be scored simultaneously for all projects as it is poorly judged by individuals familiar only with single sites. In addition, the definition of equity used here is fairly simple, and likely could be supplemented by some more meaningful empirical criteria. The value weight is not intended to reflect increased health or environmental risk associated with a site, as those impacts should be estimated in the health or environmental risk scores.
Public or Political Concern	\$600K	Relative value of eliminating highest level of concern as represented on this measurement scale. The case can be made that if community resource property transfer and environmental equity are the key reasons for public or political concern, and if these benefits are appropriately valued elsewhere in the model, then the value weight for this category should be zero or very small.

The value weights mean, for example, that the relative value of spending program or public funds is about \$5M to avert a "most

serious" statistical health impact and \$2.6M to avert destruction of 10 acres of "median" ecological habitat, both as defined in the model's measurement scales. Benefit may be expressed in terms of any value considered in the model, or an artificial numeraire of "utils," just as currencies may be translated using exchange rates.

4. An Aggregation Function

Finally, an aggregation function combines the value weights with the several technical "scores" provided via the measurement scales (and the set of single-attribute utility functions) to provide an overall measure of benefit for each project or activity evaluated.

Summary

The four concepts above form a necessary set of conceptual tools one would wish to have available when embarking on a comparative-ranking project, whether of problems or solutions. Although the formal mathematics will not be provided here for the sake of brevity, the model uses nothing more complicated than high-school algebra and basic probability, and it can be efficiently implemented in a common garden-variety spreadsheet. Multiattribute utility theory also includes methods for the assessment of value weights, rules for structuring value hierarchies, rules for defining measurement scales and rules for defining meaningful aggregation functions.

The purpose of multiattribute theory is to explain the four concepts in detail and to describe just how one proceeds to combine value-laden policy judgments with technical inputs in a comparative ranking of proposed actions. In particular, the theory tells you how to structure a value hierarchy so that your aggregation function takes on a relatively simple form, such as a linear combination of value-weighted scores on individual measurement scales. [Theorems on value hierarchies are typically of the form: Preferences involving value hierarchy, H, have property, P, involving preferences (with or without uncertainty). Aggregation function f_H using weights w_1, \dots, w_n has convenient property F_P , e.g., linear, multilinear or multiplicative.]¹⁸ The theory also makes clear just where value judgments and technical judgments

¹⁸ For details, see *Decision Analysis and Behavioral Research*, *supra* note 13, ch. 9. Comparative risk typically assumes linear aggregation, whether justified.

link up, what the desirable properties are of individual measurement scales and indeed what is meant by a measurement scale.

One of the most valuable distinctions in multiattribute utility theory, for example, is that of *means* versus *end objectives*. If, for example, you want to use pesticide toxicity as a measure of ecological hazard, recognizing that it may not also adequately represent ecological risk because neither exposure nor population size is considered. Then, you do not want to place a value weight on toxicity, because, strictly speaking, nobody cares how toxic anything is except insofar as it is a proxy for some valued outcome, e.g., ecological damage. In multiattribute utility theory, toxicity is called a "means" objective since reducing toxicity is a means to reduce what you care about, namely ecological risk. So the correct way to develop a model is to include ecological risk as the *end* objective, and recognize that you only have a proxy measurement for it. Yet, the mistake should be avoided of attaching a value weight directly to a toxicity level.¹⁹ The problem with comparative risk and some other approaches to comparative ranking is that there is no such theory guiding their development which can avoid problems of this type.²⁰

The value weights of multiattribute utility theory also have a clear meaning, as illustrated in Table 2. While some people may be repulsed by the idea of monetized tradeoffs, or tradeoffs of any kind, they should first realize that the approach here differs from traditional cost-benefit analysis in that all values are included in the definition of benefit, and that the policy weights are defined in terms of actions or choices involving the public budgets individuals wish to influence —

¹⁹ See *Value-Focused Thinking*, *supra* note 13, ch. 4 on proxy measures. Attempts to rank pesticides could be improved through a clearer treatment of proxy measures, such as LC50 toxicity values; see Leon Higley & Wendy Wintersteen, *A Novel Approach to Environmental Risk Assessment of Pesticides as a Basis for Incorporating Environmental Costs into Economic Injury Levels*, *American Entomologist*, Spring 1992, at 34.

²⁰ Another problem avoided in MUA and repeated in comparative risk is that of double-counting impacts. "Toxic pesticides" for example, may be considered as a human health and an ecological problem: so is it one problem or two? If the health impacts occur to the public and farmworkers, is that two health problems or one? The category of "problems," separated from alternatives intended to address them, leads inevitably to confusions of this type. The Colorado comparative risk report reports the problem of double-counting on its summary page; see *Colorado Environment 2000 3* (State of Colorado) (1989).

not what is "valued" by the marketplace.²¹ Value weights are explicitly and willfully adopted by a decision-making or stakeholder group through deliberation and debate; they are not the output of a survey or econometric model. The practical advantages of such an explicitly constructive approach to defining value over so-called contingent valuation approaches to benefits assessment have also been investigated by Fischhoff, Slovic and others.²² In this case, and in many other environmental policy contexts, such values represent a relative willingness-to-pay in program or public dollars and can be easily changed to show the differences, if any, that different value judgments make to the rankings or particular funding decisions.²³

For example, if you consider that the value weight on ecological impacts is too low, the weight can be changed and rankings compared for differences. While it may be difficult to articulate such judgments, they are nonetheless implicit in the decisions being made about the environment by all of us, all the time;²⁴ the multiattribute process, unlike comparative risk, simply makes them systematic and open to view by all concerned.²⁵ The approach says simply that given a set of

²¹ Value modeling as presented here contains both potentially conservative and progressive elements. A factor leading to conservatism is to neglect or mis-model synergistic or cumulative effects by evaluating single projects. Another problem is not recognizing the comparative aspect of rankings: Depending on context, even a bottom-ranked activity may be worth conducting. Progressive aspects of value-modeling are that it promotes meaningful discourse (What is environmental justice? What is an ecological impact?), makes that discourse open (in contrast to solving problems behind closed doors or in court) and requires confrontations (e.g., between economics and the environment) too often conveniently avoided by public agencies.

²² See Robin Gregory, Sarah Lichtenstein & Paul Slovic, *Valuing Environmental Resources: A Constructive Approach*, 7 *J. Risk & Uncertainty* 177 (1993); Baruch Fischhoff, *Value Elicitation: Is There Anything In There?* 45 *Am. Psychologist*, 835 (1991); Baruch Fischhoff & Lita Furby, *Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility*, 1 *J. Risk & Uncertainty* 147 (1988). For an example of how the media frame the assessment issue toward an economic-only paradigm for value, see *Poll: May Help Government Decide the Worth of Nature*, *New York Times*, Sept. 6, 1993, at 1.

²³ On sensitivity analysis, see Kristin S. Shrader-Frechette, *Risk and Rationality*, 186-187 (1991); *Decision Analysis and Behavioral Research* *supra* note 13, ch. 11.

²⁴ See Garrett Hardin, *The Tragedy of the Commons*, in *Valuing the Earth* 127, 130 (Herman Daly & Kenneth Townsend, eds. 1993); Milton Russell, *The Making of Cruel Choices*, in *Valuing Health Risks, Costs, and Benefits for Environmental Decision Making*, *supra* note 10, at 15.

²⁵ See *supra* note 17 on the "public values forum," which lawyers of some stripes

inputs and assumptions about relative importance, here is a nominal set of priorities. If you don't believe these priorities, we should be able to identify what aspect of the inputs — either technical or value-laden judgments — is the source of disagreement and proceed from there. A benefits model of this kind therefore is not to be used as a decision-maker substitute, but as an aid to dialogue and judgment; it is dynamic and open to change through the articulation and use of its measurement scales and value weights.

Basically, multiattribute utility theory provides a rigorous and consistent solution to the problems of defining and structuring values, defining measures, creating integrated rankings with competing values and incorporating project costs. The multiattribute approach to value-modeling is not as well-known as it should be; it is not perfect, but it is the most rigorous and best understood approach to modeling tradeoffs involving multiple outcomes. Uncertainty may be incorporated as well, but the greater motivation is that the key drivers of many environmental decisions are competing values rather than (or in addition to) uncertainties, and these values need to be articulated and made explicit. Following are some other issues associated with applying multiattribute utility theory that resolve continuing problems found in the applications of comparative risk.

Identifying Value Judgments

There are two main types of value judgments in multiattribute utility theory. First, a measurement scale for a single value category, such as health impacts, or ecological impacts, will itself have embedded in it value judgments. For example, if we estimate the number of acute fatalities, chronic fatalities (e.g., due to cancer) and minor injuries (made precise in whatever way) attributable to some problem, say a hazardous waste site, then, all these impacts are either valued the same, or, more likely, differently. Differences are reflected in the single-attribute utility functions mentioned above, that represent value tradeoffs "within" a single category of value. Similarly, different ecological or socioeconomic impacts may be valued differently, but *before* such impacts are compared to others, such as health or socioeconomic. Basically, even if you only have one category of value, may find unappealing for its open and non-adversarial approach to articulating values.

value judgments still relate to various types of outcomes within it. Yet, while there are no value-free measurement scales, where a value appears on a given scale is essentially a technical or "expert" judgment.

The second type of value judgment is represented by the relative tradeoffs or weights between distinct values, as shown in Table 2. While value judgments of the first type might be provided by a group of subject-matter experts, such as ecologists for environmental impacts, those of the second type might come from broader stakeholder groups. These two types are implicitly sorted out fairly well in at least some comparative risk projects; but life would be simpler knowing that there's a rigorous framework that has dealt with the issue already. Also, neither type of value judgments is called "importance," a term used repeatedly in comparative risk projects that confounds the following: A problem is important because we need to act immediately; great uncertainty is associated with it; the problem has been long neglected in the past (e.g., ecological impacts); it is the biggest problem; it will be the most costly future problem; we do not know how to solve it; or there are inequities in addressing it. *The casual use of "importance" allows all those meanings to be assumed at one time or another without pinning down the intended meaning.*²⁶ "Importance" is a holistic term for almost any factor entering into a comparative ranking and a project participant evaluating an "importance" takes on too great a cognitive burden in the absence of a structured and logical method. "Importance" is just a bad heuristic tool in the complex context of comparatively ranking either problems or solutions.

Codifying Judgment

Any comparative risk report, or almost any application of multiattribute utility theory, shows overwhelmingly a tendency to systematize expert judgment, whoever the "experts" may be.²⁷

²⁶ On the difference between "importance" and value weights, see *Value-Focused Thinking*, *supra* note 11, at 147-148 (the section titled "The Most Common Critical Mistake"); see also *Reducing Risk*, *supra* note 2, at 17 and *Environment 1991* *supra* note 4, at 15.

²⁷ On the provisional status of "experts," see *Improving Risk Communication*, *supra* note 9, at 270-271.

Table 3
Severity of Potential Ecological Impact Constructed Scale

1	No or negligible threat to resources. Exposure to the hazard source will produce no ecological effects of concern.
2	Low threat to resources. Impacts are self-correcting or of minor ecological concern, as supported by observations or knowledge of conditions similar to the following: <ul style="list-style-type: none"> • The number of species' members killed from acute or chronic impacts are expected to fall within the range of normal temporal or spatial variations. • No significant long-term adverse effects on species health are expected to occur, such as maladaptive changes in birth weight, mature size, life span, biological integrity or reproductive success. • The carrying capacity, biological diversity, ecological sustainability or abundance associated with the resource are not expected to be permanently affected at significant levels. • The likelihood that expected changes can be mitigated by human intervention is very high.
3	Moderate threat to resources. Impacts are not self-correcting and are of some ecological concern, as supported by observations or knowledge of conditions similar to the following: <ul style="list-style-type: none"> • The number of species' members killed from acute or chronic impacts may fall near the limits of normal temporal or spatial variations. • Isolated long-term adverse effects on species health may occur, such as maladaptive changes in birth weight, mature size, life span, biological integrity or reproductive success. • The carrying capacity, biological diversity, ecological sustainability or abundance associated with the resource may be threatened. • The likelihood that expected changes can be mitigated by human intervention is small.
4	High threat to resources. Impacts are not self-correcting and are of clear ecological concern, as supported by observations or knowledge of conditions similar to the following: <ul style="list-style-type: none"> • The number of species' members killed from acute or chronic impacts may fall outside the limits of normal temporal or spatial variations. • Widespread long-term adverse effects on species health may occur, such as maladaptive changes in birth weight, mature size, life span, biological integrity or reproductive success. • The carrying capacity, biological diversity, ecological sustainability or abundance associated with the resource may be significantly and permanently lowered or damaged. • The expected changes cannot be mitigated by human intervention.

Ecological risks and socioeconomic impacts often depend for their evaluation on some individual's or group's best judgment, as no model or systematized data may exist. How then, for example, should the so-

called Technical Review Groups (TRGs) found in the comparative risk project process develop useful measurement scales when it is known that no formal risk assessment will be available?

Here too, considerable groundwork has been laid in multiattribute utility theory, especially through the notion of a constructed scale.²⁸ Table 3²⁹ shows such a *constructed scale*, used as one of the inputs to score ecological impacts for the model presented already. Bullet items in the scale indicate several non-exclusive ways of providing a score--the idea being that the "experts" select the scores that best fit their knowledge of the site and its environmental condition, both "with" and "without" conducting a specific project. Those who complain that such a measurement scale is not quantitative enough should provide a substitute. Meanwhile, constructed scales provide means to represent the best knowledge appropriate to the ranking task, whatever it may be. Constructed scales are eminently suited to public participation, as they force on participants an improved clarity and explicitness about "risk," while keeping all inputs in public view. Comparative risk projects contain many scales of varying quality, but again, appear to have been developed without support of an existing body of relevant knowledge. In the multiattribute approach, unlike comparative risk, the goal of providing a complete "audit trail" back to technical inputs based on constructed or other measures is fully possible.

Distinguishing Alternatives and Values

Decision analysis is a normative theory involving three concepts: uncertainty regarding what you know or don't know, values associated

²⁸ On constructed scales, see *Value-Focused Thinking*, *supra* note 13, ch. 4. For other examples, see Miley W. Merkhofer & Ralph L. Keeney, *A Multiattribute Utility Analysis of Alternative Sites for the Disposal of Nuclear Waste*, 7 *Risk Anal.* 173 (1987).

²⁹ Note that in Table 3 "Low threat," for example, is implicitly defined by the detail following and "1-4" have no meaning in themselves. Such labels must be translated by the underlying multiattribute model into a meaningful risk calculation (making intermediate scores such as "1.5" acceptable); hence a "1" on this scale has no necessary relation to a "1" on any other scale. A common mistake in many ranking models is calculating with labels, such as using the 1-4 scale here to define arbitrary "points." But what if you change to five points, or three? How do you control your results? Note that while the judgment being asked for here is not trivial, it is one of several of a minimal set of questions one would likely want to ask when many millions of project dollars are at stake.

with outcomes of interest (the topic of this paper) and the options or alternatives from which you must choose. Environmental justice is a value and economic impacts refers to one; pollution prevention, education and risk communication refer to alternatives; reduction of consumption levels is also an alternative.

Neither environmental justice nor pollution prevention need to be elevated to "alternative risk paradigms," as appears in some comparative risk literature.³⁰ This serves only to further obfuscate and marginalize the roles of these important and interesting ideas. *How* one should value environmental justice and *how* one should measure it are difficult indeed; but there is no problem about just what its role can be in a comparative analysis. Or, to cite again the Vermont project, the authors banished "overconsumption of Vermont's natural resources" to "non-problem" status, apparently because its *bona fides* as an environmental "problem" could not be agreed upon. Insofar as a comparative risk project is intended to clarify public perceptions, the role of reduced consumption, as perhaps the only significant alternative left to Vermonters as a means of reducing their environmental problems, is thus unclear and poorly communicated by the project. Yet there should be no mystery here; "reduction of Vermont's overall consumption levels" is a decision alternative, a choice or an action, pure and simple. Similarly, *new alternatives* may be created easily for an analysis as well, such as considering site cleanups based on different land use standards or focused site studies. This philosophy of "value-focused thinking,"³¹ leads stakeholders to "work backwards" from the values of interest to new alternatives going beyond the status quo. Such thinking appears to have been part of EPA's original conception of comparative risk, but is largely missing from its current formulation.

A Limitation

A natural question is whether one can "integrate" two or more distinct sets of value weights. For example, one might want to assess the

³⁰ Some of the pseudo-controversial issues raised in the California report by the Environmental Justice Committee and Economic Perspective Committee could be addressed through a multiple-values, ranking-solutions approach even without a fully-fledged MUA model; see, e.g. *Toward the 21st Century: Planning for the Protection of California's Environment* (State of California) (1994) at 233-384.

³¹ See *Value-Focused Thinking*, *supra* note 13.

value weights through a survey or a vote, in order to define the values of a community or stakeholder group. A famous theorem proved by Nobel Laureate Kenneth Arrow says that generally can't be done.³² Unless everyone already agrees on the same set of values, a putative aggregation or integration procedure for value weights will generally violate some reasonable condition that you'd like the aggregation to have, such as being "non-dictatorial," or not being determined by any one person's values. Arrow formulated his theorem in the context of voting rules, but it applies quite readily to multiattribute value weights: any set of weights must represent either the values of a single decision maker or the consensus values of a group, but such a consensus must typically be arrived at through the usual political and social means; there is no mathematical trick to aggregate values. Nonetheless, multiattribute utility theory provides a clear conception of the relationships among "pure" technical judgments, value-laden technical judgments codified in measurement scales, tradeoffs among competing objectives — and an "exogenous" political, social or institutional process required to arrive at the value weights appropriate for a given decision-making context. Properly applied, the approach shows just where your values lie and where they come from.

Why Hasn't Value-Modeling Been Used In Comparative Risk?

First, multiattribute models can be complex. An "off-the-shelf" approach — that comparative risk projects sometimes appear to adopt — usually fails, and, because multiattribute utility theory forces all the right questions, models may become overly detailed. However, that is a matter of streamlining and knowing where to stop an analysis; any approach will have to balance accuracy, time and resources. At the same time, the decisions are often complex, include many options, reflect many concerns and may involve stakes on the order of \$100M-\$1B — perhaps annually; so a simple priority-setting model may not be useful.

The second reason even informal value-modeling may have been heretofore ignored is that the current formulation of comparative risk fits in well with EPA's attempts to replace the value-laden policy questions with quasi-technical or "scientific" solutions. As EPA has, in

³² See Ralph L. Keeney & Howard Raiffa, *Decisions with Multiple Objectives* 523 ff (1993) and Kenneth Arrow, *Social Choice and Individual Values* (2d ed. 1963).

the past, attempted to define levels of acceptable risk through science alone, comparative risk "asks the wrong question"³³ of ranking problems only or of trying to "integrate risks," whereas a methodologically sound approach ranks solutions with explicitly valued benefits. Environmental issues and methodologies to address them are inextricably linked, and EPA's scientific approach serves to limit, in advance, its influence.

A brief comparison may again indicate the extent to which comparative risk crystallizes EPA's methodological value judgment of trying to make policy by "scientific" proxies. For many years, it tried to value risks or risk reduction through methods developed by economists known as "contingent valuation," as has been eschewed in comparative risk.³⁴ Contingent valuation represents many approaches, but the central idea is that some pre-existing value is held by individuals for, say, natural resources or risk reduction; the purpose of a contingent valuation survey is to accurately estimate it. But in applications of multiattribute utility theory, values are not presumed to exist *a priori*. They may need to be articulated and explicitly constructed through public participation, risk communication or other such interventions — including comparative risk projects. A good deal of the heuristic power found in multiattribute utility theory comes from its role in helping to *define and re-define problems*, — by respecifying value categories, introducing new value categories, defining measurement criteria or introducing new risk reduction alternatives. That heuristic power, absent in comparative risk, has been termed the *constructive* aspect of the multiattribute process, in the sense that values, measures, alternatives, problems and solutions are considered to be as much "made" as "found" by participants in a multiattribute ranking.

This constructive content is part of the appeal of the multiattribute process for those interested in democratizing environmental policy debates. But the idea of constructed values, or constructed preferences, is antithetical to any approach, such as contingent valuation, which

³³ See Marc K. Landy, Marc J. Roberts, and Stephen R. Thomas, *The Environmental Protection Agency: Asking the Wrong Questions* (1990).

³⁴ See *Reducing Risk*, *supra* note 2, at 25; *Environment 1991*, *supra* note 4, at 14. On contingent valuation, see Robert C. Mitchell & Richard T. Carson, *Using Surveys to Value Public Goods: The Contingent Valuation Method* (1989).

assumes that values exist largely independently of the interventions of an assessor. In multiattribute applications, there is often only a form of analytic participant-description, through which values come to be represented in a formal model. Any complex risk ranking will require the combination of normative and descriptive methods, but multiattribute theory is simply open about that fact. Aside from its normative attractiveness, the constructive perspective has empirical support as well. One of the conclusions of cognitive research on contingent valuation techniques is that answers given to simple preference questions provided in different formats (direct assessment versus tradeoff, gambles versus certain choices, etc.) can fluctuate wildly: they are very sensitive to subtleties in how valuation questions are posed.³⁵ The answer provided by MUA is not to substitute a single new question, but rather a whole process of interaction, thinking and dialogue, through which values are expressed in improved languages of risk. So though comparative risk has not taken up the methods of contingent valuation, insofar as the problems and solutions of comparative risk lack a true constructive dimension, comparative risk will never be the type of flexible, open-ended process needed for legitimate public participation and useful policy; it will still be asking the wrong questions by not moving forward from the problem of ranking problems.

Science Begins and Ends with Problems

A mathematical science of value exists, but it is yet to be taken up by EPA, perhaps because it also leads away from science, if with clarity and method. Nor will ranking solutions using multiattribute utility theory solve all priority-setting problems. It helps considerably, but here are two examples to show that multiattribute or decision-analytic approaches also present important problems.

Incentives

Assume one had shown the limited value of a low-ranking project, but that resources potentially saved by not conducting were to be unavailable to stakeholders in any meaningful way. Would they have an interest in implementing it? Probably not; it's possible that any

³⁵ See *supra* note 22.

diversion of funds could be interpreted as a political loss if nothing differentiates that from a budget cut. For rankings to be acted upon with true self-interest, priority setting should be combined with a proper incentive structure, say car-marking saved funds for more desirable environmental protection. Such incentive schemes do not follow obviously from MUA; they must be developed and implemented in addition to a ranking model.

Mission-Driven Budgeting

How should we implement priorities? You can't just tell project managers, for example, to "Stop Project A and Start Project B;" that's centralized and dogmatic management at its worst. No approach can succeed without detailed coordination of policy-makers, stakeholders and those "in the trenches" conducting environmental work.

But the multiattribute approach is exactly that of "mission-driven budgeting" of the type needed to reinvent government.³⁶ That is, multiattribute utility analysis will link a program or organization's environmental, social and budgetary "mission objectives" to detailed line activities in the context of government and stakeholder interests. And this mission-based, "reinventing" approach is equally useful for environmental groups, regulatory agencies or industrial concerns. It doesn't matter who you are: There is a basis for communication and the improvement of environmental decision quality over time. The process consists of developing a value hierarchy with stakeholders, measurement scales with experts, sets of value weights with decision-makers and stakeholders, and inputs to score or rank alternatives, as well as evaluating the quality of inputs and seeing what it all implies. Perhaps surprisingly, while formal multiattribute utility theory can be 90% defining the *problem* and 10% the *answer*; the primary benefit often lies in improved understanding and communication.

³⁶ On mission-based budgeting, see David Osborne & Ted Gaebler, *Reinventing Government* ch. 4-5 (1993). An important consideration is the difference between measurement of mission objectives through a MUA model and a program's or business's *performance* measurement via, e.g., the number of hazardous waste sites cleaned up per year, resources expended per site or number of permits issued with pollution-prevention clauses. Performance and mission measures are related, but they are far from identical and are implemented quite differently.

Conclusion

Tools derived from multiattribute utility and decision analysis rigorously and methodically address issues posed by comparative risk. They should be considered as *means* to help define processes, methods and outputs of comparative risk projects. If nothing else, the efficiency and integrity of such projects may be improved just by making people aware of basic vocabulary, concepts and useful techniques. Much is to be gained by understanding the conceptual framework of multiattribute theory, even if all of its mathematical techniques are not implemented. Yet, the problems and contexts for environmental policy are sufficiently complex that no single approach is going to resolve them all. Progress is possible only by trying an approach and coming up with something better. The first steps are to model values, rank multiple-value solutions, not just problems, and proceed from there. Again, as the late philosopher of science, Sir Karl Popper, said: "Science begins and ends with problems."³⁷



³⁷ See *supra* note 1.