redd
3-24-86

OVERVIEW OF DECISION-AIDING METHODOLOGY

The Department of Energy (DOE) is using a formal decision-analysis method to help determine which sites appear to warrant the investment in characterizing them. Like all formal methods, this one is capable of providing only a partial and approximate accounting of the many factors important to the site-recommendation decision. Its primary usefulness is to explicate some of the key differences among the sites and to explore the significance of these differences. Due to the inadequacies of this or of any methodology, its results will not form the sole basis for DOE's siting decisions.

The methodology is relatively straightforward, although some of the terminology may be unfamiliar to most readers. It requires the identification of the most relevant attributes or objectives for siting a repository, which may be readily derived from the DOE's siting guidelines.* Judgments are then made to determine the degree to which these objectives are achieved. By doing this systematically, and by making judgments quantitative whenever possible, DOE believes decisions can be made on a more objective basis than would otherwise be possible. A more-detailed overview of the methodology in nontechnical terms follows. This overview necessarily does not cover all of the details of the methodology as applied to the siting problem, especially with regard to the postclosure analysis. It is instead intended to be a brief procedural guide.

---

*The siting guidelines (10 CFR Part 960) are organized into three categories: implementation, postclosure guidelines, and preclosure guidelines. The implementation guidelines govern the application of all other guidelines in the evaluation of sites and establish general rules to be followed during siting. The postclosure guidelines deal with the siting considerations that are most important for ensuring the long-term protection of the health and safety of the public. The preclosure guidelines deal with the siting considerations important to the operation of a repository before it is closed, such as protecting the public and repository workers from exposures to radiation, protecting the quality of the environment, mitigating socioeconomic impacts, and the ease and cost of repository construction and operation. Both the postclosure and preclosure guidelines are divided into system and technical guidelines. System guidelines contain broad repository-performance requirements that are largely derived from applicable EPA and NRC regulations. The technical guidelines specify requirements on one or more elements of the repository system. The reader is referred to the guidelines (49 FR 47714, December 6, 1984) for a more detailed discussion of their basis and structure.

To: WWalker

This discussion closely follows the development in, and at times takes sections almost verbatim from the monograph _Multiattribute Evaluation_*. The interested reader is referred to this reference and the technical literature for additional discussion and for real and hypothetical examples of applications of the methodology.

The technical name for the decision-aiding methodology is multiattribute utility analysis (MUA). The MUA technique has an extensive literature, and has been applied to a wide variety of problems, such as evaluating competing bids for various kinds of military hardware, choosing among alternative sites for businesses, and formulating positions in international negotiations. It consists of six basic steps:

1. identifying objectives for selecting among candidate repository sites;

2. developing measures to show how well each site meets siting objectives;

3. assessing numerically the performance of each site with respect to each measure, and putting these numbers on a common desirability or utility scale;

4. assessing weights for each objective;

5. aggregating utilities and weights into a composite score for each site using an aggregation rule;

6. performing sensitivity analyses.

Each of these steps is reviewed in more detail in the following pages.

The methodology is carried out by DOE staff and consultants consisting of experts in decision analysis, in the disciplines corresponding to the technical siting guidelines, and in repository performance. The technical information for the analysis is obtained from the final environmental assessments (EAs) or from references cited therein. Value tradeoffs and other policy judgments necessary for the analysis are provided by DOE management. The National Academy of Sciences' Board on Radioactive Waste Management has reviewed both the methodology and its application to the siting problem here.

---

*Edwards, W. and J.R. Newman. 1982. _Multiattribute Evaluation_. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-026. Beverly Hills and London: Sage Pubns.

Step 1: Identify and Organize Objectives

A basic premise of the decision-aiding methodology is that the "goodness" or more technically, the utility, of a site is related to the extent to which that site achieves the various objectives of site selection. Thus, the first step in the application of the methodology is to explicitly define siting objectives. It is convenient to organize the objectives in a tree or hierarchical structure, as shown in Figure 1.

The objectives at the top of the figure (such as "minimize adverse preclosure impacts") are too broad to be of practical value in distinguishing among sites. Therefore, more detailed lower-level objectives necessary for meeting the top-level objectives are identified. These lower-level objectives are easier to quantify than the top-level objectives and guide the specification of performance measures (see below).

It is necessary to indicate in the statement of the objective the direction of favorability, thus the term "minimize". It must be recognized that no site simultaneously meets all objectives to the fullest extent possible. The objectives compete in the sense that doing well against any one may mean doing worse on another (e.g., minimizing health effects versus minimizing costs). Thus, performance against some objectives must be traded off against others.

Any objectives hierarchy should capture collectively all of the important "things to consider" in making a given decision. The objectives hierarchy shown in Figure 1 is assumed to satisfy this goal because the objectives are derived from the system and technical guidelines, which were developed through an extensive process of consultation, public comment, and NRC concurrence. Care must be taken in developing the objectives hierarchy to avoid double counting objectives. Extra or unnecessary objectives only serve to make the analysis more complex.

Step 2. Develop Measures to Quantify How Well Sites Meet Objectives

Having developed a hierarchy of objectives, the second step in the decision-aiding methodology is to develop "yardsticks" to indicate how well a site meets objectives. Formally, these yardsticks are known as performance measures. The development of performance measures is, in our experience, the most difficult of the steps in the methodology; it is essentially a creative process that requires professional judgment, knowledge, and experience. Ideally, performance measures are objective and based on physical measurements or hard data. Inevitably, however, some measures are "intangibles" that are not easily described or quantified.

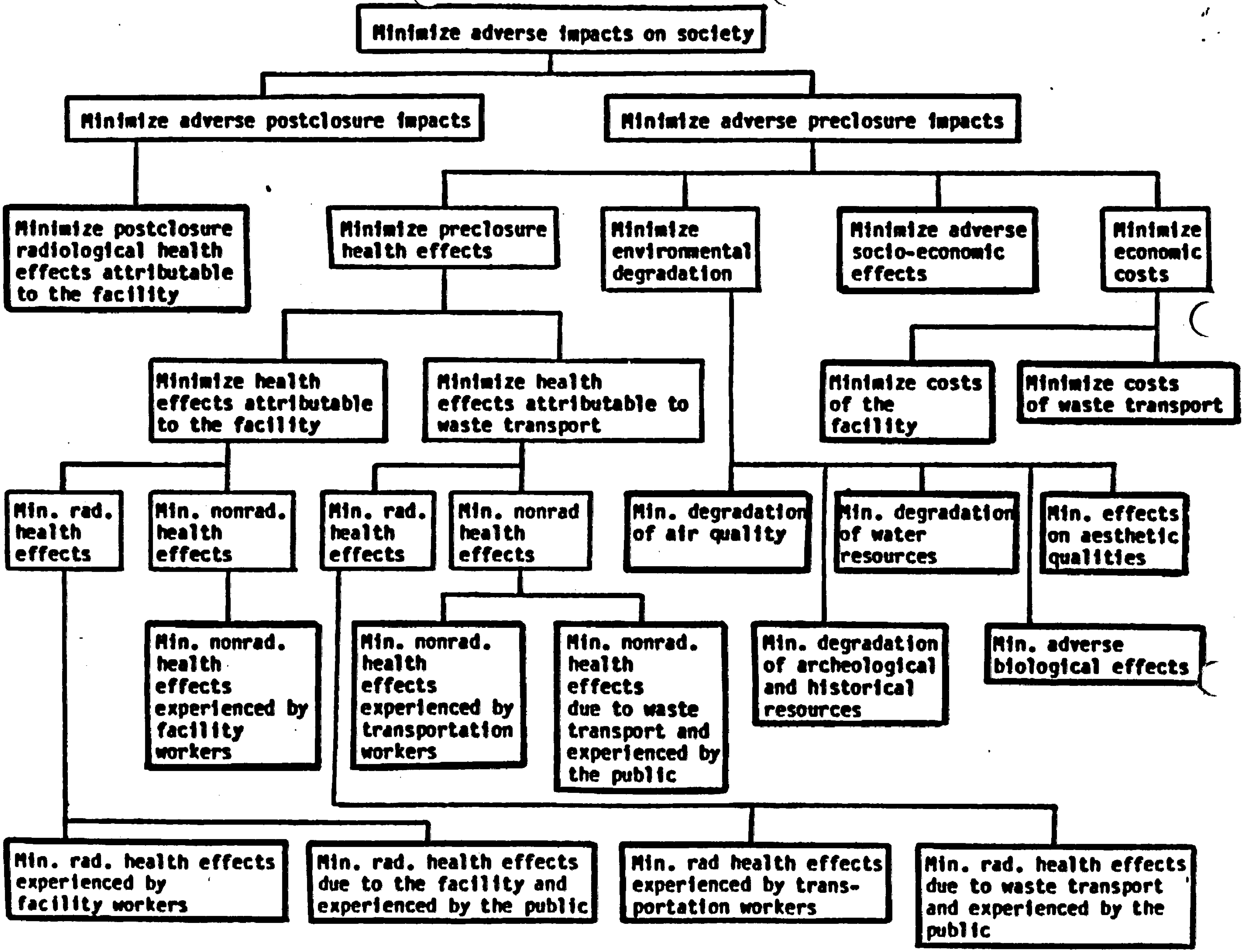

Figure 1. Objectives Hierarchy for Repository Siting

Hypothetical examples of performance measures are shown in Figures 2 and 3. As may be seen, 0 to 10 scales are defined for each measure in terms of detailed site descriptors (Figure 2) or more natural measures such as dollars (Figure 3). Although the choice of the scale is arbitrary, we choose one in which 0 means a level of performance for a very poor site and a 10 means a level of performance for an ideal site. The ranges spanned by performance measures should be realistic in the sense that it should be imaginable for sites to score 0 or 10 on each measure. Where possible, the ranges of the scales are defined without specific reference to the performance of the actual sites being evaluated, i.e., the scale is absolute. This is preferable because, if new sites turn up, they too are likely to fall within the range. However, for some measures, because of the difficulty of setting reasonable upper and lower bounds (e.g., costs), the ranges are chosen with reference to the actual sites available, i.e., the scale is relative. It should be noted that the definitions of the 0 and 10 have no effect on the final rankings computed in a multiattribute utility analysis. (The upper and lower bounds do affect the value of the weights, however, as explained later.)

One additional point concerning Figures 2 and 3. Total facility cost is an objective measure, measurable in dollars. Environmental impacts is somewhat more subjective; the measurement of it relies more on judgment. Such mixes of objective and subjective measures are commonplace in applications of the MUA, and are largely unavoidable whether or not MUA is used. A strength of the MUA technique is that it makes such judgmental evaluations explicit and produces an audit trail.

In this particular application of the MUA technique, a graphic device (not shown here) known as an influence diagram is constructed for each performance measure. Influence diagrams show the factors and the interrelationships among these factors that must be accounted for in scoring the sites. The most important factors in the influence diagrams are the technical site descriptors ("right hand side") in the performance measure.

Step 3: Score each site on each measure, and put numbers on a common scale.

The next step in the methodology is to make detailed assessments of each of the sites (on the basis of data in the EAs) with respect to each performance measure. Such assessments result in a numerical score between 0 and 10 for each site for

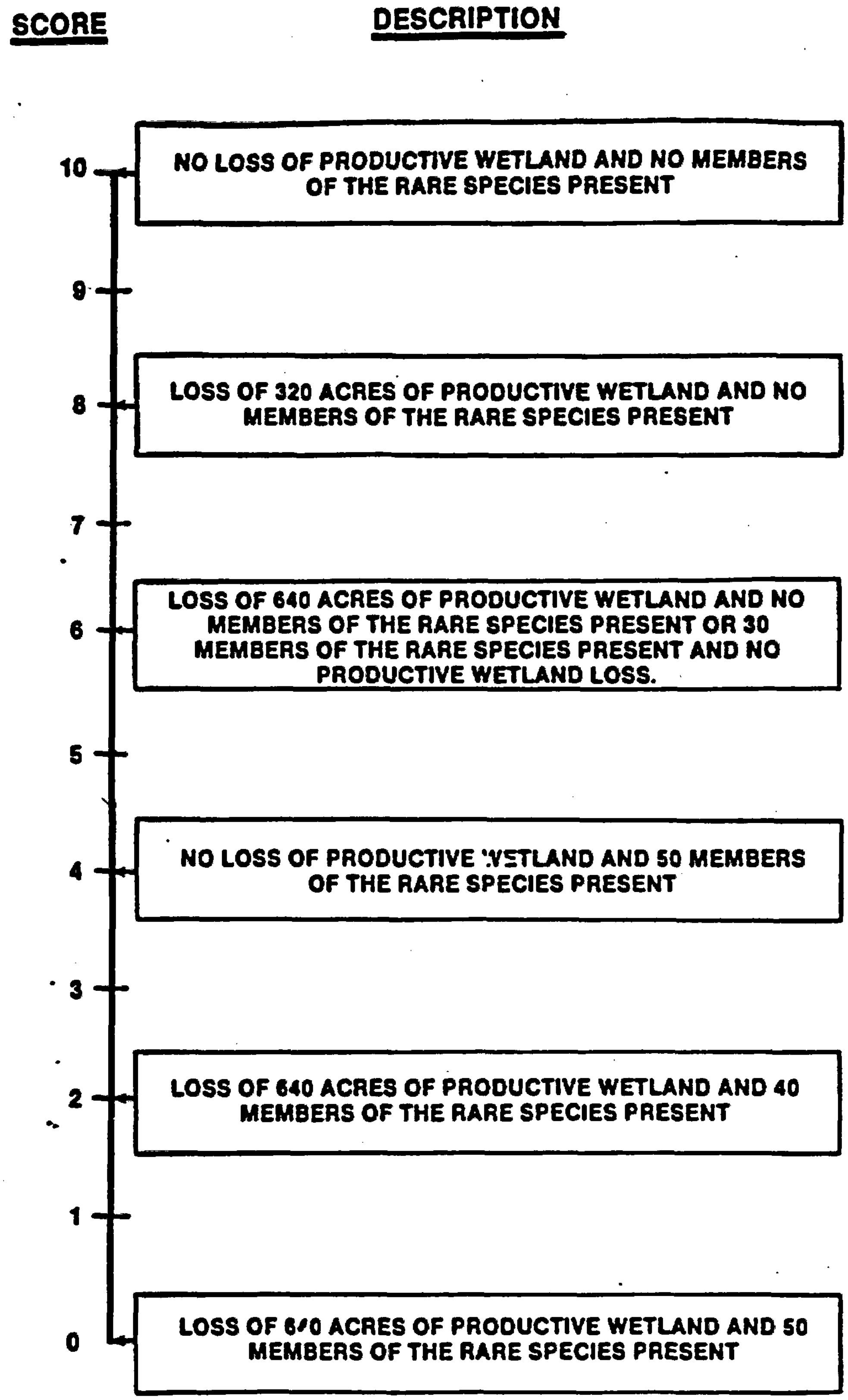

Figure 2. Hypothetical performance measure for objective "minimize impacts to rare species". Modified after Keeney, R.L., 1980. Siting Energy Facilities, Academic Press, New York.

0213-0023MP 1/16/86

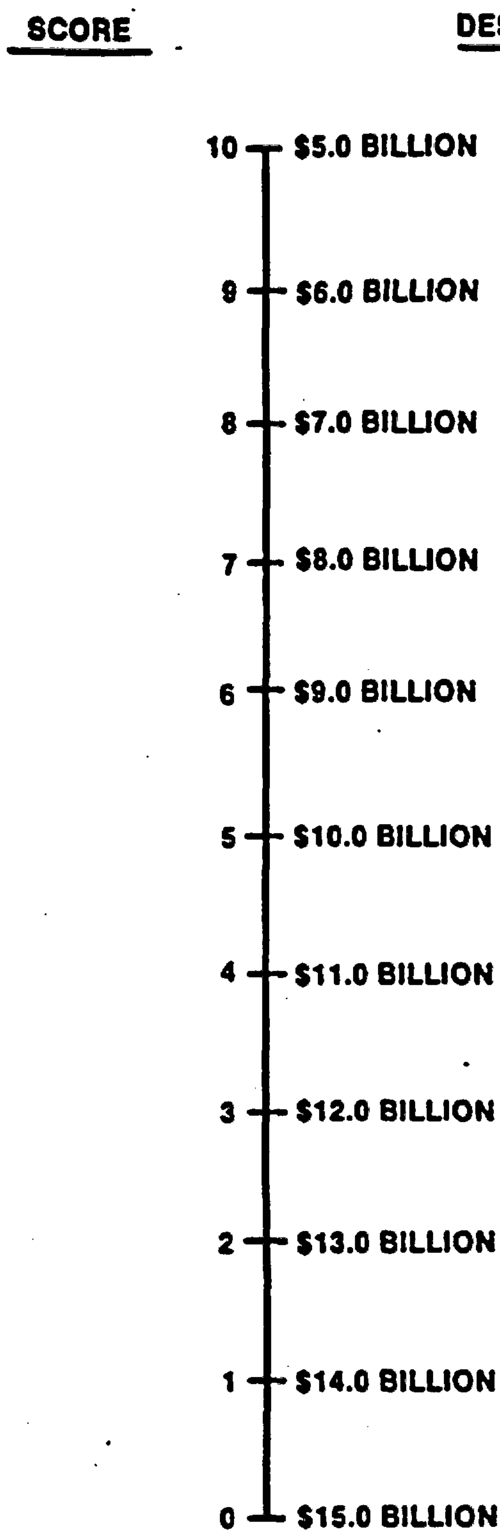

Figure 3. Hypothetical example of performance measure for objective "minimize total facility cost."

0213-0023MP/16/86

each of the lowest-level objectives. These judgments are made by a panel of individuals selected for their expertise on the topic of the assessment.

To account for uncertainty, three scores will be assigned to each site for each performance measure. First, a score representing the most likely performance level will be assigned. Second, a minimum score will be assigned to indicate the worst performance judged plausible. Finally, a maximum score will be assigned to indicate the best performance judged plausible. It is possible that the uncertainties may be larger than the differences among the sites.

Such raw numbers are not enough, however. They must be transformed to a common desirability or utility scale to permit the assessments of weights (Step 4) and to account for differences in the importance of various scores. Again, the choice of a common scale is arbitrary. We use here a scale of 0 to 100 in which 0 means horrible and 100 means as well as one could hope to do.

The transformation of the raw scores (0 to 10 scale) to the desirability or utility scale (0 to 100 scale) is illustrated in Figure 4 with references to the two hypothetical performance measures shown previously. The horizontal axis of the graph is in units of the particular performance measure of interest (as represented by the 0 to 10 scores), here in units of environmental impacts or dollars. The vertical axis goes from 0 to 100, and is in units of desirability, technically called utiles. To assign utility, simply locate the score on the horizontal axis and read off the corresponding utility on the vertical axis.

As may be seen from the graph, each increment of performance on the 0 to 10 scale does not necessarily have equal value. For example, the intent of the environmental objective might be mostly met by a site that scores only a 5 on the 0 to 10 scale. Such a site might receive an 80 on the utility scale. In other words, improvements in performance beyond a 5 might bring only marginal returns in terms of desirability. Such a relationship between desirability and a physical measure is termed nonlinear, and would plot as a curved line. This is analogous to the law of diminishing returns in the field of economics. The more common situation is for desirability to uniformly increase or decrease with the physical measure over the whole range, here illustrated for costs. The relationship would be termed linear, and would plot as a straight line as shown.

One utility curve is derived for each of the performance measures. Because such judgments are largely policy judgments, not technical, they are elicited from the policy makers at the DOE.

Step 4: Weight Each Objective

The remaining problem to be overcome in the analysis is that the lower-level objectives themselves are not equally important in attaining the overall objective of minimizing repository impacts. Weights are developed to account for this. The decision-analysis literature has devoted much attention to defining weights and to weighting procedures. In general, the weights must be established to reflect the relative importance of a swing from the lowest to highest score against an objective, not some ill-defined generic importance of the objective. Thus, weights reflect a policy makers's willingness to make value tradeoffs among performance measures. The weights are intimately tied to the performance measures; if the ranges of such measures change, then the weights must also change.

Weights are value judgments. As such, they are likely to vary from person to person. It is for these reasons that weights are varied in a sensitivity analysis (Step 6). It often happens that the final result is relatively insensitive to different weights.

Step 5: Aggregate weights and utilities into a composite score using an aggregation rule

At this point in the methodology, two sets of numbers are available: weights, one for each objective, usually calculated so as to sum to 1.0; and utilities calculated for each site for each of the objectives expressed on a scale from 0 to 100. The next step is to aggregate these two sets of numbers into a composite score using an aggregation rule. The simplest and most commonly used rule is a linear additive one. This rule involves taking the weight for each objective, multiplying it by the utility a site achieves for that objective, and summing the products over all the objectives.

This deceptively simple aggregation rule embodies a number of subtle and important assumptions. The most important assumptions are concerned with the independence of the various objectives. If, for example, it may be shown (by a formal process not described here) that the importance of doing well on one objective does not depend on the level of performance achieved on any other, then one can often assume that the weights for specific objectives are constant across the various sites, certainly a desirable feature of the analysis. If dependencies exist, then the analysis becomes more complicated. This is why the individual technical guidelines cannot be used as siting objectives, and simply scored, weighted, and added. Great care is taken in the analysis of the siting problem here to ensure

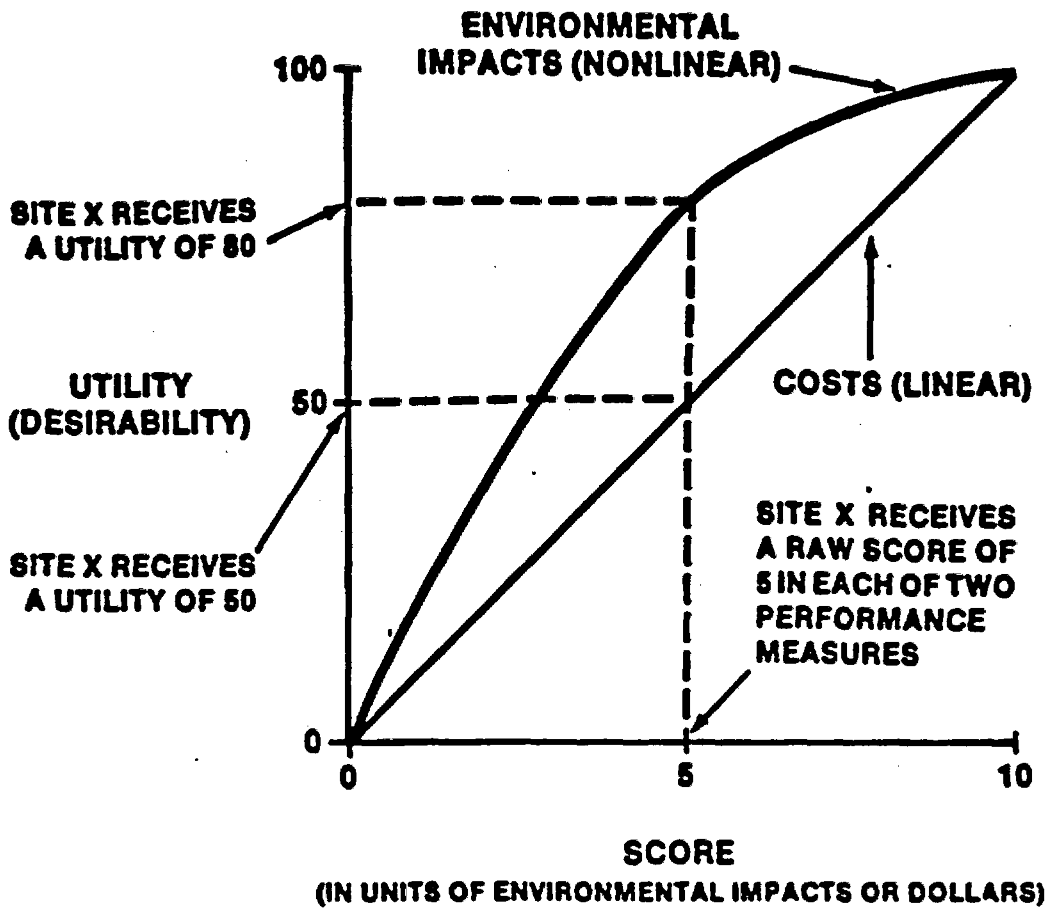

FIGURE 4. Examples of graphs relating utility (desirability) to physical measures. Modified after Edwards and Newman, 1982.

that these underlying assumptions hold so that the simple aggregation rule can be used.

Step 6: Perform sensitivity analyses

The purpose of sensitivity analyses is to test how the original overall utilities, calculated in step 5, change as assumptions and judgments change. If the conclusions from the original analysis are resilient under changes in assumptions and judgments, they are more likely to be valid. An obvious sensitivity analysis is to vary the weights, since different people have different opinions on the relative importance of the various siting objectives. Other input data to the methodology, such as the scores (step 3), may also be varied.

In summary, one of the major assets of the decision-aiding methodology is that it breaks the problem of selecting sites for characterization into its component parts, which can then be analyzed more easily and scrutinized more readily. The methodology does not reduce the subjectivity or professional judgment required in selecting sites for characterization. By following the sequence of steps outlined above, however, DOE hopes to make the subjectivity inherent in the scientific and policy judgments explicit to the reviewer. The methodology does this in essentially five ways. First, it specifies and organizes DOE's siting objectives. Second, it provides a means for summarizing how well each site meets each objective. Third, it provides a means for specifying alternative judgments about the relative importance of each objective. Fourth, it provides a systematic way to "add up" sites' scores on individual objectives. Finally, the methodology allows DOE to test how conclusions change as judgments and assumptions change.