Abstract
A man-machine interactive system, Kanshiki, is designed to augment and sharpen man's judgment in evaluating alternative designs or plans of action. Both factual and value-oriented information are needed, and can be treated consistently and systematically, in evaluating possible consequences of proposed alternatives. The adequacy of criteria used can be questioned and their definition and structure modified, through interaction, to explore many trade-off implications and judgmental interpretations. With Kanshiki, man will be able to examine a much larger number of alternatives, weighing many different factors, than he normally can.

Introduction
The major concern in this paper is how to evaluate programs with value-laden issues such as those that are directed toward improving the "quality of life"—any programs, in fact, having socioeconomic, ecological, political, or psychological implications. We all know that technology has given us mixed blessings and that we are paying a high price for some of the "benefits" gained. Because technology often intermeshes intimately with the social fabric, we are becoming more and more aware of the necessity for long-range planning and for assessing possible consequences of our actions. It is in this planning and problem-solving context that I am going to discuss a technique of evaluation.

To place the evaluative process in a proper setting, I shall first present a simple characterization of the four decision steps commonly found in planning and problem solving. They are:

- Define the objectives and set appropriate criteria.
- Generate alternative courses of action.
- Identify or estimate possible consequences of each alternative.
- Evaluate the consequences in terms of the criteria and choose the alternative which best achieves the objectives.

Since the initial attempt at defining objectives and criteria is often inadequate and incomplete, these steps are usually repeated. The iterative nature of the process is shown schematically in Figure 1.

![Figure 1. Iterative Decision Steps](image-url)
Difficulty in evaluating alternatives. Problem situations that are complex and ill defined often defy conventional cost/benefit analyses in evaluating alternative courses of action. Major difficulties are discussed below.

• There are many important aspects of life that escape objective measurements. Definitions of "costs" and "benefits" can be expressed in terms of "undesirable aspects" and "desirable aspects," respectively, but they are seldom well defined, much less quantified.

• Even if they are made explicit and quantified, a single "quantitative measure of effectiveness" is seldom adequate to summarize the issues that arise because of the variety of impacts.

• Differences in individual value orientation cause both obvious and subtle divergences in judgment-oriented decision making. Furthermore, such value systems do change over time and in different circumstances, even for the same individual.

• Trade-off implications are complex and confusing even to an experienced decision maker because of the many factors involved and their incommensurable values. Intuitive judgment is seldom reliable in such a situation. When the complexity of the situation exceeds the capacity of man to cope with it, oversimplification and premature conclusion often result (see reference to "cognitive economy" in Hormann, 1971, Part I).

• There is a growing evidence that the likelihood of making "knowledgable" decisions decreases rapidly as number of alternatives and number of criteria are increased (see the typical curves shown in Figure 2 below).

This seems to be true even with (or because of) a large body of available information. Today's decision maker is often overwhelmed with many pieces of information, each of which may be relevant but not conclusive in itself.

In addition, what the decision maker usually has or receives is a mixture of factual and value-oriented information, qualitative and quantitative measures, and objective and subjective judgments, some of which come to him as "raw" information but many are generated, transformed, and inferred during pre-evaluation and evaluation periods. With these, decision makers face the task of evaluating a large number of alternatives, weighing many factors of both a tangible and intangible nature.

Summary of the approach taken and the rationale. My approach to this type of evaluation task is to use Kanshiki, the man-machine interactive system, which includes a set of techniques and programs as a tool incorporating the "fuzzy-set" concept (Zadeh, 1965). Use of this tool is the major topic of the discussion. A description of the "fuzzy-set" concept will be given later; its meaning suggested by its name will be sufficient now to permit one to understand the rationale behind our approach.

• Although the value-laden issues will never be "solved," i.e., complete agreement on all the issues or universal acceptance of proposed program or design is not possible, the problems are here today and must be dealt with. We must act on the current information available with best techniques available now.

• The approach proposed here is an attempt to bridge the gap between analytical procedures and intuitive procedures; they should be used in a complementary fashion. For this purpose, existing analytical methods and decision aids, such as those developed in utility and decision theories, have been incorporated into Kanshiki to lift many burdens from the decision makers, but more reliance will be placed on judgments and intuition (1), which are elicited by special interactive techniques.

Since the emphasis is on systematic analyses supplemented by intuitive judgment, we strive toward consistency and logical comparability in making relative judgments by providing explicit treatment of indeterminacy and various degrees of imprecision. It may be far more serious to omit a criterion that is believed to be important, just because it cannot be made precise, than to include it at a low level of accuracy.
Man-machine interactive facilities and techniques, including interactive visual displays, can provide immediate feedback and flexible means of testing and adjusting criteria for evaluation, taking into account intangibles that are often excluded in analytical methods. However, man’s judgment of alternatives typically becomes less reliable when multiple criteria of varying importance must be concurrently considered. Kanshiki can "evaluate" many alternatives rapidly once criteria for evaluation and methods of aggregating individual evaluation in various aspects are specified. In the man-machine context, man is the specifier of criteria and rules, which he can change as he makes different interpretations of previous results, and as he examines trade-off implications in the light of new findings.

Immediate feedback and the ease of adjusting previous value assignments allow evaluations to be made serially and in context, as new insights are gained. Thus the traditional methods of having to decide in advance all the factors to be considered and all the conventions to be followed can be relaxed considerably. This will encourage man to explore many "what if" questions and trade-off implications. Man will be able to examine a much larger number of alternatives, weighing many different factors, than he normally can before a final decision or selection must be made.

Kanshiki can provide a means of assessing and weighing different "points of view" or value orientations in group evaluation. Systematic methods, such as the Delphi technique, can be used for the direct involvement of experts and users in exploring possible consequences, priority issues, trade-offs, and even policy changes.

Differing opinions and preferences are mostly due to differences in individual value orientations (and institutional or "public" values) and in degrees of understanding various aspects of a situation. Group interaction through Kanshiki can influence subsequent rounds of evaluation, not by group pressure, but by reasoning and increased understanding.

Polarization may occur, but a completely even distribution (of the evaluators’ value inputs) is not likely for most value-laden questions. Experiments have shown that individuals tend to be more responsive to value-oriented questions than to factual questions; that is, changes of opinion seem more readily attainable on value-oriented issues than on factual ones. Usually, no new values are introduced, but restructuring of values seems to take place in the light of new understanding about many sides of an issue.

Qualitative or descriptive information in evaluation, intermixed with quantified information, will eventually be reduced to numbers using Kanshiki. However, using a conceptually "natural" method of eliciting a number (value assignment) from a person is as important as choosing appropriate aggregation rules for integrating these values.

The "fuzzy-set" concept and associated techniques are used to provide explicit treatment of "matter-of-degree" judgments (e.g., "how safe," "how comfortable," and "how convenient"), and to provide relative judgmental situations (rather than forcing absolute judgments).

An Example of Evaluation

Many areas of application are possible and I can describe Kanshiki in completely general terms, but an example will be helpful in making the presentation clearer. Suppose a group of commissioners (evaluators) are evaluating a number of proposed plans for developing a park (local, state, or national). There are many criteria for determining the desirability of such plans, but for brevity let us say that "peaceful atmosphere" and "utility to the public" are the two most important ones. Since these are very general, such criteria are usually described in terms of component attributes such as "number of acres of vegetation or foliage," "number of feet or miles of streams," "variety of flowers," for the first one; and other attributes such as "number of picnic tables," "number of benches," "number and sizes of parking lots," for the second one. The latter group is relatively easy to evaluate for appropriateness since a set of desirable standards (per acre of park) is usually known. But the general criterion of "peaceful atmosphere" is much harder to represent since the total effect cannot be perceived readily from the components of the first group.

Scale models of proposed plans are often useful for visual overall evaluation, but it would be prohibitively expensive and time consuming to provide models for all the alternative designs, $A_1$, $A_2$, ..., $A_n$, submitted (2).

One useful tool is a visual input/output display scope connected to a computer for man-machine interactive use (3).

Here, the terrain characteristics of the park-to-be area are displayed (preferably in color) to the group of evaluators and then a proposed design of the park is displayed with an option of enlarging any part of it for display (4). Some verbal descriptions and numerical information such as cost and those attributes mentioned earlier (number of picnic tables, variety of plants and flowers, etc.) can be added.
Each evaluator is now asked to give his opinion of the design $A_i$ in relation to the criterion, 'peaceful atmosphere.' This is expressed as "grade of membership" (in the set of all alternatives for which the criterion applies) in terms of a number in the interval [0, 1]. If the number is close to 1, say 0.94, then $A_i$ has a high grade of membership as far as this attribute is concerned; if it is close to 0, say 0.1, it is not a highly valued member. The question may be phrased in the following manner: 'If 'peaceful atmosphere' is the only criterion with which to judge the proposed designs $A_1, A_2, ..., A_n$, and if all the designs can be graded to take relative positions on a straight line between 0 and 1, where would you place the design $A_i$ compared to your highest ideal that would be at the position 1?'

Since comparability is important in qualitative judgment, other alternative designs from $A_1, A_2, ..., A_n$ can be displayed one or two at a time for comparison. Two or more A's can take on the same value in case of a tie. The evaluator can change his mind about the values he previously chose. The "grade of membership" he gives for the first design alternative he considers may be more or less arbitrary, but as he proceeds in the comparing process, the values tend to indicate the relative merits of the proposed designs.

All evaluators in the group make their evaluations independently of each other, so the set of values collected may differ greatly. Group interaction with or without anonymity and re-evaluation of the set of values can be the subsequent step, using the on-line or off-line Delphi technique (5) (more about this will be discussed later), or the evaluators can proceed to another attribute, still independently of each other. The only thing they must agree on first is the initial set of attributes. These can be changed later, but the group must agree on the changes.

Fuzzy-set concept. I have tried to set a stage for an intuitive understanding of the "fuzzy-set" concept in the above discussion. Let us now clarify the notion of "grade of membership" mathematically. Suppose $X$ is the set of all alternatives. Let $A$ be a subset of $X$ for which attribute $a$ applies. In our park example, $a$ may be "peaceful atmosphere." A person's subjective preference judgment can be represented by a preference relation which associates $A_i$ with a number in the interval [0, 1]. To this relation, I have assigned the symbol $R$. The value $R(A_i)$ represents the "grade of membership" of $A_i$ in set $A$ (see Figure 3).

![Figure 3. Fuzzy-Set Association](image)

The two values $R_1$ and $R_2$ show a possible difference in two persons' judgments about $A_i$ on its grade of membership for attribute $a$. This tends to make more visible subtle individual differences in value orientation.

So far, we have been talking about types of attributes for which there is no sharp transition from membership to nonmembership. Of course, there are other attributes that can be defined in a non-fuzzy fashion, i.e., a "yes" or "no" answer (1 or 0 value) can be given; then its membership function becomes identical with the characteristic function of a non-fuzzy set.

Trade-off considerations. One of the most important benefits we get from the use of "grade of membership" is that trade-off concepts can now be dealt with quantitatively. Suppose in our example of the park design, "utility to the public" has also been graded for each alternative. One factor of utility to the public may be "accessibility to many parts of the park by car." But many people feel strongly that this requirement will be in conflict with "peaceful atmosphere." It is true, but exclusive concentration on "peaceful atmosphere" and little accessibility will deprive some segments of the public (e.g., those who are infirm) from enjoying the total facility. Then, trade-off implications must be explored. Questions such as "How much 'peaceful atmosphere' can be traded for how much 'public utility'"?, which were meaningless in conventional evaluation, can now be treated sensibly because the two attributes are now comparable, represented by the same unit of measurement "grades of membership" (see Figure 4).

Even if all attributes were quantified (e.g., number of picnic tables, number of flowering plants, etc.), they are still incommensurable and the trade-off concept does not apply. It is, therefore, important to evaluate grades of membership for all the attributes, even though some attributes are naturally quantifiable, such as...
"cost of developing the park" and "cost of maintenance." Here, costs are still evaluated on a comparative basis.

The essence of this approach can be stated as follows: If it is inappropriate to quantify everything and reduce the measures to a single "measure of effectiveness" (e.g., dollars), then change everything into value-oriented judgment.

The essence of this approach can be stated as follows: If it is to quantify everything and reduce measures to a single "measure of effectiveness" (e.g., dollars), then change everything into value-oriented judgment.

Even after all the attributes have been considered for all the alternatives, those values and vertical lines, though comparable, do not constitute the making of an overall judgment. One way to facilitate it is to calculate the summation

\[ S_j = \sum_{i=1}^{m} w_i \cdot R_a(A_j) \] for all the weighted "grade of membership" values over \( m \) attributes for each alternative \( A_j \), where \( w_i \) is a weighting specification to indicate the relative contribution of \( a_i \) to the achievement of goals and objectives (ranking of attributes and assigning weights will be discussed later) (7).

\[ S_j = \sum_{i=1}^{m} w_i \cdot R_a(A_j) \]

\[ \frac{S_j}{\sum_{i=1}^{m} w_i} \] (called the summary value) for each \( A_j \) can be done by Kanshiki very rapidly. Now comparing summary values \((S' \text{'s})\) is a meaningful process since their values are again in the interval \([0, 1]\), representing overall grades of membership.

**Individual differences and group interaction.** Since each evaluator will have his own set of summary values \((S' \text{'s})\), there may be collected diverse values of \( S' \) for the group of evaluators. This divergence can come from differences in \( R_a \) values and/or \( w \) (weights) values. The Delphi technique or simpler group interaction may be used at this time. The Delphi technique may have been used earlier for both \( R_a \) and \( w \) values, but the evaluators are not likely to come to a complete agreement on one set of values (but eventually there will be fewer sets than the number of evaluators). Any group interaction may yield some influence toward agreement. Some studies on group planning seem to indicate that people may widely disagree on objectives and criteria but may readily agree to favor certain alternatives.

Let us see how different sets of \( S' \text{'s} \) for different evaluators can be compared. Let \( S' \) represent a composite attribute of all \( a_i \) 's that have been considered. In our park example, \( S' \) will be "acceptability of a park design with all the features and weights properly taken into account." Summary values can now be used as \( R_a(A_j) \)'s and a display of \( R_a \)'s for all alternatives \( A_1, A_2, \ldots, A_n \) can be made for each evaluator (see Figure 5). The total display of all such values of all the evaluators may be shown in a scrambled order to maintain anonymity, if desired. In addition, a statistical group response, such as quartiles \((Q_1, M \text{ (median), } Q_3)\) of each \( R_a(A_j) \) value, can be calculated. Seeing where his own evaluation stands within the group response may aid him in understanding the overall evaluation.
As an evaluator studies the relative merits of alternatives and gains a deeper understanding of trade-off implications, his total conception of the situation may become more defined. If $R_B(t)$ is his preference function operating on the fuzzy-set at time $t$ with respect to the global criterion imposed by $B$, then as $t$ increases, $R_B(t)$ tends to converge toward a more precise function; i.e., the evaluator becomes better able to sort out alternatives.

Imprecisionness of many decision-making situations is usually caused by a mixture of ignorance, randomness, and intrinsic fuzziness. This kind of exercise can help the evaluator to separate out types of imprecision involved and learn to identify (a) where more information is needed (case of ignorance), (b) where probabilistic treatment is needed (randomness), and (c) where matter-of-degree judgments are needed (intrinsic fuzziness) that require increased awareness of other value systems besides his own.

In complex decision situations where many competing factors must be properly accounted for simultaneously, the interactive system can be made to keep track of the evaluator's tendencies. For example, suppose the evaluator is excessively cost oriented and his assignments of grades of membership for the cost attributes fall consistently outside the interquartile range $(Q_1, Q_3)$ of the group response. The system can remind him of other important factors and trade-off considerations.

Interacting with other evaluators through the system and seeing where his own evaluation stands within the group response may influence him to take additional factors into account or to adjust his preference functions. If he feels strongly about his preferences, he can try to persuade others by stating the reason why the value should be lower (or higher) than the values (in the interquartile range) expressed by the 75 per cent majority.

The process of interaction and reevaluation can be repeated until, it is to be hoped, convergence is attained. Polarization may occur but completely even distribution is not likely for most value-laden questions. Individuals seem more responsive to value-oriented questions than to factual questions; that is, changes of opinion seem more readily attainable on value-oriented issues than on the factual ones, especially when relative impacts that certain factors make on the overall effects are realized by group discussion.

Steps of Machine-Aided Evaluation
The foregoing discussion in the park design example introduced only the fuzzy-set concept and the use of the Delphi technique. In this section, more nearly complete steps of machine-aided evaluation are described. It should be noted, however, many refinements that can be added through the use of utility theory, Bayes' theorem, simulation, sensitivity analysis, etc., are not covered here because of the obvious space limitation. Those concepts and techniques that are useful in the man-machine context are emphasized and others deliberately oversimplified.

There are seven evaluation steps to be followed, some of which will be elaborated on in the following pages. For simplicity, the procedure assumes a single user (evaluator) and group interaction is not emphasized. Steps are presented in the "usual" order but can be reordered at the user's direction.

1. List alternatives by name or number assigned to each.
2. List criteria for evaluation in terms of attributes.
3. Rank attributes and assign weights.

4. List values in their "natural" description (numerical or non-numeric) for each alternative's attributes.

5. Determine grades of membership of all values of attributes.

6. Calculate the summary value of each alternative.

7. Repeat any or all the steps above.

The list of alternatives and attributes may be prepared in advance covering steps 1, 2, and 4, and can be thought of as an attribute-alternative table (Figure 6). Unlike mathematical tables, this table can contain both numerical and non-numerical descriptions, even lengthy discussions supplemented by pictures that can be referenced. Therefore, the physical form of the information may not look like the table in Figure 6.

Figure 6. An Attribute-Alternative Table

In the following, expanded descriptions of the seven evaluation steps are given where necessary.

1. List alternatives by descriptive names or by distinct numbers assigned to them.

2. List criteria for evaluation in terms of attributes.

Attributes can be given on a noncommittal trial basis with full recognition that they are likely to be inadequate or incomplete; or they can be carefully selected by a group of people (e.g., policy makers, planners, experts, representatives of the public). Attributes may be separated into two groups, "desirable attributes" and "undesirable attributes", or they can be all mixed together. Subsequent instructions will reflect the choice.

At this stage, threshold values of certain attributes can be specified in terms of lowest and/or greatest values that are acceptable (e.g., the lowest acceptable load capacity and the highest acceptable cost for a vehicle). This allows some alternatives that do not meet the crucial requirements to be eliminated before proceeding with detailed evaluation. Threshold specification also aids the decision makers to distinguish between what are essential and what are desired.

3. Rank attributes and assign weights.

Attributes are rank ordered in terms of their relative importance in contributing to the objectives. If there are many attributes and ranking is difficult, Kanshiki can assist the evaluator by presenting only two attributes at a time. Judging the relative importance of two attributes is much easier than ranking the whole list. However, the resulting ordered list should be reexamined as a whole to guard against any possible context shifts, which may result when only two items are compared at a time.

If the evaluator's judgment of importance is transitive and total in ordering, the attributes are listed and displayed in the order of importance, possibly placing two or more in the same rank in case of a tie. The logic (computer programs) inside Kanshiki can easily check any inconsistency and ask the evaluator to compare again those attributes whose rankings are in conflict. If the inconsistency is not removed, it is likely that at least one attribute should be redefined to eliminate overlapping membership with other attributes. Other redefinitions often useful are substituting two or more other attributes in place of the current one and grouping some attributes together as one. An experienced evaluator usually can sense which ones are in need of adjustment (8).

When the ranking is complete, the attributes are displayed in rank order. The evaluator is now asked to assign weight (w) to each attribute, starting at the top-ranked one(s) with the weight of, say 100. Using this as the point of reference, the other attributes are also assigned weights. These weights should reflect the relative "strengths of effects" of attributes contributing to the objectives.

4. List attribute values in their "natural" description (numeric or non-numeric) for each alternative's attributes.

If some attribute values are originally given in verbose descriptions, they can be condensed to a few key words to be displayed along with the
names of attributes. The original information sheets should also be available to the evaluator. Those quantified attributes, for which threshold values have been specified, will be consulted by Kanshiki at this stage for preliminary sifting out of alternatives.

5. Determine "grades of membership" of all the attribute values.

Using the fuzzy-set concept, each attribute value is judged in terms of "grade of membership"—i.e., a number in the interval [0, 1]. Since comparability is important in value judgment, other values of the same attribute from different alternatives can be shown one or two at a time for comparison. In Figure 6, this process corresponds to moving horizontally across the alternative on the same attribute line. When all the attribute values are judged in this way, a new table of values is created within the uniform scale. It will then look exactly like Figure 6 containing a single number (between 0 and 1) in each box in the table. When these values are displayed in bar graphs (see Figure 4), visual comparisons can be made easily and the evaluator, seeing the total picture of value assignments in terms of bar heights, may wish to change his earlier choice of values. A simple light-pen action on the display scope will lengthen or shorten a bar to his specification. Trade-off possibilities can be explored at this time.

6. Calculate the summary value of each alternative.

The summation \( S_j = \sum_{i=1}^{m} w_i R(A_i) \) and the summary value \( S_j' = \sum_{i=1}^{m} w'_i(1 - R(A_i)) \) are calculated for each alternative \( A_i \) as one of the basic machine aids, but some other forms of getting the summary value may be tried out. The evaluator can specify his own ideas easily with the man-machine communication language, User Adaptive Language (UAL) (see Hormann, et al., 1970).

If attributes have been separated into "desirable attributes" and "undesirable attributes," \( S_j \) is calculated using only those \( \alpha_i \)'s in the desirable category, and \( S_j' = \sum_{i=1}^{m} w'_i(1 - R(A_i)) \) is calculated using only those \( \alpha_i \)'s in the undesirable category. The difference in their summary values, \( S_j/w_i - S'_j/w'_i \), where \( w_i \)'s and \( w'_i \)'s are weights attached to those attributes for \( S_j \) and \( S'_j \) respectively, may be called the "net-benefit value" for each alternative \( A_i \). Comparing these values will result in a tentative conclusion concerning which alternative(s) is(are) best.

7. Repeat any or all the steps above.

The evaluator is encouraged to go back and examine his previous judgments. It is usually advisable, the first time around, to use first impressions in making attribute rankings and in making a judgment of grade of membership without too much deliberation. Stepping through the whole sequence rather quickly the first time, rather than dwelling on a single factor in detail, will give him a better understanding of how certain factors are accounted for in the total evaluation.

Iterating the evaluation process tends to bring many assumptions into the open, and the evaluator may become more aware of how the conclusions are related to the assumptions. For example, assumption on the objectives will influence the interpretation of objectives and criteria in terms of attributes and will also influence attribute ranking and weight assignments. Assumptions on political and technical constraints on the proposed designs will certainly influence many decisions. Probing into them with "what if" questions may separate out "real" constraints from imagined ones or those that can be overcome by negotiation or by creative problem solving.

The evaluator may, in the light of new insights and understanding, wish to redefine objectives and specify relevant attributes more carefully. Interacting with the other evaluators, or even with the policy makers, may bring further clarification. Possible use of the Delphi technique has been discussed; it can be used repeatedly in any of those steps described, but it will be most useful after step 6.

In studying complex trade-off implications that are typically nebulous, evaluators might gain a new perspective by asking Kanshiki to display an additional bar graph besides the one shown in Figure 4. Bars in Figure 4 can be rearranged to show the \( R \) values of different alternatives horizontally for each \( \alpha_i \) (Figure 7).
For group interaction using the Delphi technique, Kanshiki can provide another display, showing all the bar graphs generated by different evaluators for each attribute \( \alpha \) (similar to Figure 5 but for one \( \alpha \) at a time, not for the overall \( \beta \)). The evaluators may ask "what if" questions on different value interpretations and trade-offs (either collectively or individually) that are not evident in the design alternatives; the answers may suggest a new or an improved design.

Potential Applications
There are a number of potential applications for Kanshiki, including evaluation of architectural designs and urban development plans. They are categorized in four groups in terms of their characteristics:

1. **Complex equipment with many performance criteria.**
   
   Evaluation of different designs of complex equipment such as aircraft and underwater exploration vehicles can use machine-aided evaluation. In these, many attributes must be included in evaluation and they cover both factual information and subjective value information. This class of problems is less fuzzy not only because factual information tends to dominate but because the physical boundary in measuring operational behavior is relatively clear.

2. **Selection of suitable locations for large complexes.**
   
   The problem of selecting a suitable location for a large complex, such as a model city development, often requires careful consideration of many attributes that are qualitative in nature. Among many possible locations, one or a few candidates are usually selected in order to proceed with designing, legal and financial negotiations, etc. Other examples of large undertakings whose location decisions tend to affect various segments of our society are: manufacturing plants, airports, hospitals, health-care centers, sanitoriums, rehabilitation centers, educational institutions, trade centers, highways, and transportation networks.

3. **Complex combination of things that interact.**
   
   Making an appropriate EDP system selection from all possible combinations of available hardware/software products to meet the user needs is a complex problem. Guessing at a suitable hardware/software mix is hard enough, but evaluation of a wide number of configurations when the components interact usually requires an advanced modeling technique (Sutherland, 1971). Information on the performance characteristics of hardware and software components are separately available, but very little information can be had on the total performance characteristics for specific configurations—unless the pieces are all of the same manufacture. After modeling produces the system's performance characteristics, our technique can be used in total performance evaluation.

   A similar situation facing the decision maker is the selection of alternative designs of hospitals, universities, housing complexes, or research laboratories.

   Another area of interest is compensation programs to provide employees many different options. Companies who can provide many options will have a definite advantage in employee inducement and retention.

4. **Long-range large scale programs comprised of many projects that are interrelated and interdependent.**
   
   Many government programs such as health care, welfare, education, foreign aid programs, and other research and development programs are in this category. This is an area of great importance because of its far-reaching effects, both intended and unintended. It is also the area of greatest difficulty because of its complexity, unclear boundary (sphere of effects are not clearly definable) and future-oriented consideration.

   These programs or measures that tend to create many side effects or that produce long-term effects or irreversible conditions, must receive extra care in planning. Although the future is always uncertain and, therefore, no forecasting techniques can claim total accuracy, a variety of forecasting techniques combined with modeling can produce some indication of types of impacts a given program might make in the future.
After possible consequences of alternative courses of action are generated, the consequences can be arranged within a "decision-event map," indicating interrelation of actions taken, their intended results and possible side effects, and intervening events that are likely to happen.

Concentrating on the consequences in the time-stream (rather than at one point in time), our technique, modified to include subjective probability estimates, can still be employed, using attributes that indicate future impacts (e.g., "rate of yearly increase in food production in country X, during 1970-1975, after introduction of farming equipment" or "number of farm workers in country X migrating yearly into cities during 1970-1975").

Admittedly, any future-oriented evaluation is very tenuous. However, evaluating proposed U.S. programs to assist underdeveloped countries is a more amenable problem than evaluating our own future possibilities. We can use the U.S. and other developed countries as models in planning to avoid possible undesirable consequences and to promote those attributes that are desired by the country. Although exact correspondence between the model and the real consequences in a given country cannot be expected, hindsight is readily available while foresight is not.

Summary
The importance of including many criteria of various types and degrees of imprecision has been discussed. The man-machine evaluation approach described here is our first attempt to tackle this task. Insight gained in using these techniques may lead to improvements or to new ideas and techniques.

Systematic analyses of the situation supplemented by intuitive judgment was emphasized. The following points may be worth reviewing:

- **Consistency in treatment of all alternatives with many attributes describing desirability or undesirability.** One aspect of consistency achieved here is the making of everything into a value-oriented judgment; even though attribute values may be objectively measurable, determining their worth in relations to the objectives requires a judgmental decision. The fuzzy-set concept allows explicit treatment of imprecise value judgments.

- **Comparability.** Since absolute judgment is far more difficult than relative judgment, the man-machine techniques facilitate the ease of comparison by bringing in other relevant factors in a visually comprehensive manner. In addition, the fuzzy-set treatment of attribute values make them commensurable, and complex trade-off possibilities can be explored much more readily than without such assistance.

- **Systematic use of the knowledge and experience of experts as well as opinions of people from different backgrounds.** Those techniques (such as the Delphi) for direct involvement of people can increase the acceptability of the evaluative decisions, if not the quality of evaluation.

Notes

1. This reliance permeates every aspect of the systems approach—in circumscribing the extent of the system; in deciding what hypotheses are likely to be fruitful; in making appropriate interpretation of policies, objectives, and constraints; in making assumptions about cause-and-effect relations and about information-gathering and processing requirements; and in putting facts together to develop a bigger picture by interpreting the results of information processing.

2. Techniques of making scale models, I am told, are being improved. They can be constructed much more quickly and less expensively than in the past. Also, there is a promising field of holography used in viewing three-dimensional structures that have been designed but that do not exist physically yet. Holographic plates are artificially synthesized from the given design and the known characteristics of interference patterns with respect to the structural characteristics.

3. See INTUVAL (Kamnitzer and Hoffman, 1970). This work concerns interactive design in urban planning, but the same facility and the technique can be extended to assist evaluators who may not be professional designers. It may be desirable to have a mixed group of specialist-designers, government officials, representatives of diverse civic groups, etc., to avoid personal biases as much as possible.

4. Ideally, each evaluator should be provided with an interactive facility. The evaluators work independently but can interact with each other through the system. To approach realism in evaluation, an existing park of similar size and purpose (if possible, one that nearly everyone likes) may be chosen. Photographs of this actual park and a display of how it looks in abstraction will facilitate understanding of how a proposed park will be expected to look.

5. See Dalkey (1969) and Helmer (1966). The technique was initially used for soliciting and collating experts’ opinions in long-range forecasting, but many experiments have been conducted with members of the public for estimating desirability of certain programs, and understanding people’s attitudes and value orientations. The major characteristics of the technique are anonymity, iteration and controlled feedback, and statistical group response.
6. In Figure 4, costs are shown in their grades of membership as the "preference" measure. Therefore, the lower cost, which is usually preferred, is rated high (the longer vertical line), and the higher cost is rated low (the shorter vertical line); this may be counterintuitive. The evaluators may prefer to group together as "cost" all the attributes that should be minimized (e.g., pollution and noise) and as "benefits" all the attributes that should be maximized. Then these two groups are displayed separately (side-by-side but grouped together) and new values, $R' = 1 - R$, will be used for "cost" attributes (then a high cost shows a high vertical line).

7. This form of getting a single value for each alternative is not the only way. Some non-linear or nonuniform way may be used to account for different nature of attribute definitions.

8. Hierarchical structuring of attributes from general to specific is often useful in reducing inconsistent ranking, when many attributes are being considered in evaluation. Ranking is done within each level and recursive use of Kanshiki is possible for evaluation of each level separately, while automatically keeping track of other levels of evaluation and aggregation.

References


