MULTI-CRITERION DECISION PROCEDURES AND THE ASSESSMENT OF POLICE DEPARTMENT PERFORMANCE*

John M. Greiner
The Urban Institute
Washington, D.C.

Introduction

Public administrators are increasingly recognizing the importance of basing decisions on multiple performance criteria. It is widely acknowledged that a single criterion--for instance, cost--is usually inadequate as a guide for decisionmaking in the public sector. Public services are, in fact, inherently multidimensional: they involve numerous goals and objectives, each of which is often itself characterized by several performance measures.[1] Consequently, public administrators must, in principal, examine and reconcile many different criteria when considering questions of agency performance.

Advances in public management and decisionmaking over the last two decades have intensified the need for dealing with multiple measures of effectiveness. The growing availability of practical quantitative measures of the efficiency and effectiveness of public services has made it more difficult for public administrators to continue to focus on only one or two performance criteria. Moreover, the increasing use of procedures such as management by objectives, program budgeting, and advanced capital budgeting techniques has compounded the need for considering--and combining--multiple measures of performance.

* This report is based on research supported by Grant No. 82-IJ-CX-0032 from the National Institute of Justice. The author would like to thank Harry Hatry and Dr. Annie Millar of The Urban Institute for their assistance with this study.

Indeed, it appears that the more systematic the approach to performance measurement and public management, the greater the diversity of the information needed to adequately characterize and manage agency performance. For instance, two recent "systems approaches" for assessing organizational performance--Total Performance Measurement (TPM) and Organizational Assessment (OA)--emphasize the need for regular readings of numerous agency attributes and performance indicators. Thus, TPM stresses the importance of monitoring employee attitudes and job satisfaction (from employee surveys); agency (and worker) efficiency, effectiveness, and service quality (from "hard" measures of agency performance); as well as customer satisfaction (from client or user surveys).[2] OA demands an even greater variety of information, ranging from measures of organizational efficiency and effectiveness to characterizations of overall organizational structure, job design details, employee attitudes, and inter-unit relations.[3]

However, these and other advances in public management pose a dilemma: the better that public administrators understand public services and respond to the need for multi-criterion characterizations of agency performance, the more difficult it becomes to effectively utilize the resulting information for making decisions. Increasingly, public administrators must confront large numbers of varied performance criteria. At any given time, some of these criteria may be improving, some declining, and some holding steady. How is an administrator to untangle the resulting "rat's nest" of trend lines in order to know whether the department's overall performance is any better than it was before? Similar difficulties beset many other common decisions faced by public administrators--the evaluation of employee performance, the selection of capital budgeting projects, the adoption of alternate service delivery strategies, the selection and evaluation of productivity improvement initiatives. All, in principle, involve difficult multi-criterion decision problems.

Without an effective way to deal with multi-criterion decisions and assessments, recent progress in performance measurement and performance monitoring for public services may be seriously undermined. The value of having a comprehensive set of timely, accurate performance measures can be negated if the overall picture is too hard to discern, or if the measures must be used in conjunction with crude, imprecise decision techniques. All too often decisionmakers overwhelmed by numerous performance measures invoke ad hoc simplifications and heuristics to make the assessment or decision problem tractible. The arbitrary assumptions, approximations and implicit weightings that result can seriously undermine the quality of the final decision, despite the availability of good performance data.

Indeed, the absence of practical ways to utilize multiple performance criteria for decisionmaking purposes can contribute to management frustration with attempts to regularly monitor the performance of public services. Such efforts have too often led only to long lists of performance indicators tacked onto monthly reports or the annual budget. This failure of performance measurements to become anything more than "window dressing" can often be traced at least partly to the absence of simple, practical techniques for synthesizing the mass of available data into meaningful overall assessments of performance and regularly bringing the results to bear on important management decisions.

Multi-Criterion Assessments in Police Departments. The urgent need for ways to handle multi-criterion assessments and decisions is perhaps nowhere clearer than in the case of local police departments. Police agencies have moved farther and faster than most public services in developing and implementing comprehensive systems of performance measures.[4]

Exhibit 1 illustrates the diversity of the performance data currently--or potentially--available to police managers. (These measures were compiled from data prepared in several police departments.)

EXHIBIT 1

Types of Performance Measures Currently or Potentially Available in Municipal Police Departments

SERVICE QUALITY AND EFFECTIVENESS MEASURES

Reported crimes per 1000 population (Part I, Part II)*
Reported rate of "observable" or "suppressible" crimes*
Victimization rate (total crimes--reported and unreported--per 1000 households)*
Percentage of households victimized at least once by crime
Property loss from crime (dollars) per 1000 population (vehicles only, total)

Persons arrested (adults, juveniles, total)*
Arrest rate: arrests per 1000 reported crimes*
Quality of arrest: percentage of (i) adult arrests, (ii) felony arrests for which:

- Police withdraw charges, cancel charges, or release without charging
- District attorney refuses to file charges for police-related reasons
- Charges are dismissed at preliminary hearing for police-related reasons
- Defendant is tried and judged not guilty

Clearance rate: percentage of reported crimes cleared (by arrest, total)*
Percentage of stolen property value recovered (vehicles only, total)

Percentage of emergency/high priority calls responded to within "X" minutes
Percentage of non-emergency calls responded to within "Y" minutes
Average or median response time (by priority level)
Percentage of citizens calling police who are satisfied with response time

Percentage of citizens who feel unsafe walking in their neighborhood at night

Percentage of citizens who rate police officers as being generally (i) fair, (ii) courteous, (iii) helpful
Percentage of citizens requesting non-crime assistance who were satisfied with the service received
Percentage of citizens rating overall police crime-control services as good or excellent

Citizen complaints regarding police misbehavior (total, number and percent sustained)
Number of police officers disciplined for misbehavior

Number of traffic accidents (injury accidents, fatal accidents, total)
Number of (i) pedestrians, (ii) all persons injured in traffic accidents
Number of (i) pedestrians, (ii) all persons killed in traffic accidents
Traffic enforcement index: convictions from citations or arrests for hazardous moving violations ÷ number of fatal and personal injury accidents

*Reported in total and by type of crime.

EXHIBIT 1 (Continued)

EFFICIENCY MEASURES

Arrests per employee-year or per $1000 expended*
Arrests surviving first judicial screening per employee-year or per $1000
Arrests through investigative followup per investigator-year or per $1000 expended on investigation
Arrests through investigative followup surviving first judicial screening per investigator-year or per $1000 expended on investigation

Clearances per police employee-year or per $1000 expended*
Convictions per police employee-year or per $1000 expended*

Average cost per hour for police patrol
Average cost per case solved via investigation
Average cost per traffic accident responded to

INPUT/WORKLOAD MEASURES

Staff levels (sworn, civilian, total)
Staff hours
Calls for service answered (crime-related, total)
Patrol-miles driven
Cases investigated

EMPLOYEE ATTITUDE MEASURES

Employee job satisfaction levels (management, non-management)
Employee morale levels (management, non-management)

Incidence of counterproductive behavior:

- Separation rate (voluntary, involuntary)
- Absenteeism rate (rate of sick leave usage)
- Rate of on-the-job accidents, personal injuries

The measures in Exhibit 1 also illustrate the interactive, partly redundant nature of many public sector performance criteria. Such criteria often exhibit explicit inter-dependencies or strong statistical relationships, features that further complicate multiple criteria decisionmaking. For example, the presence of a common variable can lead to explicit interactions or interdependencies between performance criteria (for instance, between the rate of reported Part I crimes and the rate of supressible Part I crimes, or between the arrest rate [arrests per 1000 reported crimes] and arrest efficiency [arrests per employee year]). There may also be strong statistical relationships--redundancies (positive correlations) or conflicts (negative correlations). Redundancy can occur inadvertently when measures turn out to partially overlap, for instance by focusing on imperfectly differentiated aspects of a single underlying characteristic (e.g., the percentage of citizens rating the police as generally courteous and the percentage of citizens rating the police as generally helpful, or the reported crime rate and the household victimization rate). Redundancy can also be consciously built in to provide multiple perspectives or data sources, e.g., measures of actual response times (from police dispatcher records) and measures based on citizen assessments of police response times (from a citizen survey); see also the multiple quality-of-arrest measures noted in Exhibit 1. On the other hand, at any given time, certain pairs of performance measures may exhibit strongly <u>conflicting</u> trends--e.g., the number of citizen complaints regarding police misbehavior vs. the percentage of citizens who have rated police officers as being unfair.

As a result, at any given point in time, police managers are likely to face a tangle of interactive, partly redundant upward- and downward-moving performance indices. How, then, is a police chief to respond when a city councilman asks whether the agency's overall performance in controlling crime is any better--especially when (say) response times, arrest rates, and citizens' feelings of security have improved, yet crime rates, citizen complaints about police harrassment, the cost per arrest, and police morale have deteriorated? Asking an administrator to provide a single overall assessment based on such diverse information may seem somewhat simplistic. Nevertheless, such questions are asked, and justifiably so. Indeed, they are likely to become even more frequent as the availability of good performance data--and public (or media) awareness of it--increases. Moreover, similar questions arise daily within the department--e.g., in the course of assessing patrol strategies, reviewing the results of special programs, evaluating supervisory performance, justifying budget requests, deciding on capital investments, etc.

Up to now, police managers (and other public administrators) have had little guidance on how to make sense out of the multitude of changing performance measures that have become available. What appears to be needed is a practical management tool for routinely synthesizing and utilizing multiple performance criteria, whether it be to understand results obtained in connection with regular performance monitoring; to respond to questions from management, legislators, the media, and the public; or to make decisions involving several performance measures.

Formal Multi-Criterion Decision Procedures. The issues posed above represent different versions of the classical multi-criterion decision problem. This problem can be stated as follows: for a set of alternatives characterized by multiple attributes, select the one (or the set) which is, in come sense, "best". The final selection may also be subject to various types of external constraints or conditions and various levels of uncertainty. The constraints can be in the form of explicit or implicit relationships between the various performance (decision) criteria, externally specified goals or standards of performance, or "process" relationships between agency inputs, outputs, and outcomes.

If the objective is to determine whether overall agency performance has improved, the "alternatives" for the multi-criterion decision problem described above would be individual multi-attribute assessments of agency performance for various time periods. Or the "alternatives" could, instead, be an employee's (multi-criterion) performance appraisals for several years, and the problem could then be to decide whether the employee's overall performance had improved. Similar formulations can be used to assess the effects of new programs and policies: here the "alternatives" would correspond to multi-criterion assessments for agency performance before and after program (or policy) implementation.

The above formulation is not limited to problems involving longitudinal comparisions of performance. Agency managers may wish to assess the relative performance of several contemporary entities, each of which is characterized by multiple performance attributes. In this case, the "alternatives" associated with the multi-criterion decision problem might correspond to the performance of different individuals, various police teams or patrol units, proposed budget or program options, etc. Moreover, most formal procedures for resolving multi-criterion decision problems generate an index of the overall value or utility of each alternative. This index reflects all the performance criteria and--with experience--can often be used to rate overall performance without reference to specific alternatives.

In recent years, there has been a great deal of research on the topic of multiple criterion decisionmaking. This research has involved a variety of disciplines and perspectives, including management science, operations research, the cognitive sciences, statistical decision theory, and others. Over the past decade, in particular, there has been much progress in developing procedures for addressing multiple criterion decisions, and numerous techniques have emerged.

These advances hold considerable promise for helping public administrators. Up to now, however, developments in multi-criterion decisionmaking appear to have had little effect on public administration in general and police departments in particular. While there have been some public sector applications (see below), they have been rather sporadic and have focused primarily on specific problems or projects. Rarely have they been applied to the assessment of overall agency performance.

Indeed, many of the multi-criterion decision procedures (MCDP's) that have been developed are rather complex and esoteric. The literature on MCDP's has tended to be quite technical and inaccessible to public administrators. Moreover, applications of MCDP's to public sector problems have usually been initiated and executed by academics or consultants and written up in very technical form. Thus, despite their high level of development, multi-criterion decision procedures have not yet demonstrated much success in becoming an accepted tool routinely used by public managers for dealing with multiple performance criteria and related issues.

The purpose of this paper is to acquaint police administrators with the range of procedures that have been developed to deal with multi-criterion decision problems and to assess the applicability of these procedures to practical decisionmaking problems routinely faced by police officials--in particular, the assessment of overall police department performance. The primary emphasis will be on the potential of formal (mathematical) MCDP's as a practical tool for police officials to use in regularly monitoring overall performance. This paper is based on the results of a study conducted by The Urban Institute for the National Institute of Justice between June, 1983 and December, 1984.

In the remaining sections, we first provide an overview of the major types of approaches available for dealing with multi-criterion decision problems. (We have tried to emphasize qualitative descriptions of the procedures wherever possible to keep the mathematics to a minimum.) The various classes of formal MCDP's are then reviewed in terms of several characteristics important to police decisionmaking, and applications to police and other public sector problems are described. This is followed by a discussion of the requirements that formal MCDP's must satisfy if they are to be practical for routine use by police and other public sector decisionmakers in monitoring overall agency performance. On the basis of these requirements, four especially promising MCDP's are identified--social judgment theory, multi-attribute utility technology, the analytic hierarchy process and compromise programming. We conclude with an assessment of these techniques--and MCDP's in general--as tools for police decisionmaking.

Systematic Procedures for Dealing with Multi-Criterion Decisions and Assessments

For ease of exposition, the major procedures for dealing with multi-criterion decisions and assessments can be divided into three categories: unidimensional approaches, non-mathematical group techniques, and formal (mathematical) procedures.

Unidimensional Approaches. Probably the most common procedure for handling multiple decision criteria is to focus on a single performance criterion--for instance, a particular cost or benefit measure. The remaining performance criteria are then handed in any of several ways: sometimes they are ignored; sometimes they are treated as constraints (e.g., by ensuring that they at least exceed certain minimum levels, an approach known as "satisficing"); sometimes they are invoked only to break ties with respect to the primary criterion of interest (a technique known as "lexicographic ordering"); and in some instances they are translated to the same units as the primary criterion. In the latter case, the criterion of interest becomes a common denominator--the "numeraire" of classical economics. If the common denominator is cost, all performance criteria are translated into dollar equivalents, and the procedure becomes the familiar "cost-benefits analysis." The latter approach has often been deemed especially appropriate for multi-criterion program analyses and decision problems in which dollar impacts are of paramount concern, for instance in the selection of capital budgeting projects or the assessment of alternate environmental options and policies.

Unfortunately, decision approaches putting primary emphasis on a single performance criterion are usually inappropriate for assessing police and other public services. For such services, no single performance criterion clearly overshadows the others in importance. Furthermore, efforts to translate police performance criteria to a single common denominator - for instance, dollars - are usually fraught with uncertainty. How, for instance, can one put a dollar value on an arrest, a complaint of police harrassment, citizen feelings of security from crime, or the level of morale within the department? Any effort to do so usually entails so many assumptions and approximations that distortions are inevitably introduced and the results quickly lose credibility.

Non-Mathematical Group Techniques. These procedures employ structured small group processes to resolve complex questions--for instance, how to select and weight diverse performance criteria, or whether (given various performance results) overall agency performance has improved. These techniques involve few, if any, mathematical manipulations; instead they rely on the use of carefully controlled social processes to seek a concensus. Two procedures are especially prominent--the "Nominal Group Technique" and the "Delphi Method."

In the "Nominal Group Technique," a group of five to nine persons is guided by a trained facilitator through a series of steps designed to identify, explore, and reach concensus on responses to a specific question.[5] The question should be carefully formulated in advance, and the members of the group should have some expertise in the area addressed by the question. The following basic steps are involved in a nominal group session:[6]

1. After an introductory statement by the facilitator, each participant is asked to respond to the given question by (privately) jotting down his or her ideas. No discussion is allowed, and only minimal clarification of the question by the facilitator is permitted. The silent generation of ideas lasts from 4 to 8 minutes.

2. A "round robin" phase is next. Here, each member of the group is asked to concisely state one idea from his or her list. The idea is present-

ed only as a brief sentence or phrase, without discussion or elaboration; it is recorded verbatim by the facilitator on a flip chart. The facilitator continues to ask each member of the group in turn for another idea until all the ideas generated by the participants have been presented. As each flip chart page becomes filled, it is taped to the wall for all to see.

3. After the round robin presentation is completed and all ideas are listed, each item on the list is read and briefly discussed. Typically two minutes of discussion are allowed per item. Questions, comments and expressions of approval or disapproval are encouraged. However, the author of an item is not specifically responsible for clarifying or defending it. Extensive arguments or efforts to resolve conflicts are generally avoided. After all items have been considered in turn, comments may be offered on any item, and a few new items may be added to the list.

4. Next, each participant selects what he or she believes to be the five (or perhaps seven) best responses from the list of items and orders them by priority. The participant records his or her selections and priorities on 3 x 5 cards and submits them to the facilitator. The results are then tallied and discussed. The item receiving the most consistently high ranking is identified as the most preferred response to the original question.

The "Delphi Technique" also involves a small group process, but it usually avoids face-to-face debate while providing structured feedback to participants on the opinions of the other members of the group.[7] Information is exchanged through a sequence of questionnaires: after each iteration, the questionnaires are returned to a monitor who summarizes the results, identifies the overall group "position" and the degree of concensus on that position, and develops a new questionnaire based on those results. Participants are encouraged to share (in writing) their opinions and the reasons for those opinions--especially if they differ substantially from others in the group--and they are allowed to revise their opinions at each iteration of the process. The process continues until there is no further progress towards a concensus, for all practical purposes.

There are many variations of the Delphi method, including the Policy Delphi, Cross-Impact Analysis (for addressing strongly interactive issues), and others. However, the basic process can be illustrated by the problem of estimating a specific number, for instance a priority rating or utility level. The assembled group of experts would address this problem as follows:[8]

1. After the problem is described to the group, each participant is asked to provide an independent anonymous estimate of the number in question and to submit it to the monitor.

2. The monitor orders the various responses and determines the median and the interquartile range of the estimates (the interquartile range corresponds to the two estimates that delineate the middle 50 percent of the responses). This information is then communicated to each participant via a questionnaire.

3. Participants have a chance to revise their original estimates after learning the range within which most estimates fell. If a participant's original or revised estimate lies above or below the interquartile range, he is asked to explain briefly (on the questionnaire) why he believes that the figure should be lower (or higher) than the estimates made by 75 percent of the participants.

4. The monitor summarizes the results of this round by indicating the new median and interquartile range for the estimates, as well as the reported reasons for extreme estimates (the lowest and highest 25 percent). This information is communicated to the participants via a new questionnaire.

5. The participants are asked to (individually) review the new median and interquartile range, as well as any reasons or opinions provided, and to revise their estimates in the light of this information if they feel it is warranted. Any revised estimates that fall outside the (new) interquartile range must also be explained on the questionnaire.

6. The monitor once again summarizes the results, computing a new median and interquartile range for the estimates and listing any new reasons for extreme values. The results are again communicated to the participants in a questionnaire, and step 5 is repeated.

7. Steps 5 and 6 are repeated until additional iterations bring little or no change in the median or interquartile range of the estimates. The median is then used as the best estimate of the number in question.

The nominal group process and the Delphi technique can be quite useful in helping to resolve certain multi-issue problems. Both procedures have been used from time to time to set priorities among multiple criteria or alternatives and to develop other inputs needed by some of the formal MCDP's discussed later (e.g., the multi-attribute utility technology). However, such procedures appear to have some drawbacks vis-a-vis their mathematical counterparts, especially when it comes to applications involving the regular assessment and comparison of overall agency performance. In particular,

- In group decisions and assessment approaches, much of the process leading to the final outcome is hidden from view--the procedure is, in effect, a "black box." The relevant assessment an decision mechanisms--and the factors that contribute to them--are usually only dimly perceived. There is, moreover, no assurance that all of the available performance criteria have been given careful, systematic consideration.
- The implicit weights and values that the participants use in conjunction with such procedures are often not delineated. Thus, a sensitivity analysis of the results can be difficult or impossible.
- No quantitative index of the overall "value" of an alternative is produced in the course of most group procedures. Thus, it can be difficult to assess the need for clustering close results; similarly, it may be impossible to compare the results of assessments by different participants or at different times.
- If similar assessments are needed later, the group process usually has to be repeated from scratch, preferably with the same participants (or mix of participants) in order to ensure consistency.
- A group approach may not be able to serve the needs of a <u>single</u> decisionmaker (e.g., a police chief) since the technique incorporates inputs from a number of individuals, persons whose perspectives or value structures may differ markedly from that of the primary decisionmaker. Moreover, group techniques tend to blur differing perspectives. Individual contributions--and differences--are usually minimized or lost in the final synthesis, making it difficult for the primary decisionmaker to determine which perspectives are dominating the result.

- Most importantly, group approaches do not obviate the need for each individual participant to resolve for himself (at least in the first iteration) the very problem that has proven so intractible--how to reconcile and combine a large number of interactive, partly redundant performance criteria into a single overall assessment or decision. Group techniques do not facilitate these individual multi-criterion assessments, except perhaps by providing feedback on why and how others are making the same difficult judgments.
- On the other hand, group techniques do appear to force each participant to confront multi-criterion problems, to carefully reflect on the participant's own assessment of the problem, and to consider the opinions of others in making that assessment. Moreover, once individual multi-criterion assessments have been elicited from a number of persons, group approaches appear to be useful in forging a synthesis of the different perspectives represented.

Taken together, it would appear that non-mathematical group techniques will be of limited use for addressing many of the multi-criterion decision and assessment problems of interest to police management (e.g., for periodically synthesizing multiple measures of police department performance into an overall assessment of whether performance has improved). Nevertheless, group approaches can help with the estimation and/or assessment of certain specific inputs (single-criterion utility functions, criterion weights, etc.) needed in conjunction with formal MCDP's. In addition, they may be helpful for combining MCDP results reflecting the perspectives of different decisionmakers.

Formal (Mathematical) Techniques.[9] Many of the techniques developed for handling multi-criterion decisions and assessments involve the use of mathematical procedures. Mathematical techniques can be called upon to help structure or simplify the decision criteria; to reduce the number of alternatives that must be considered; to assist in the specification of criterion weights; to create a single index of value incorporating some or all of the criteria; to search among given alternatives for those deemed "best" (or to guide the decisionmaker in searching for the best alternative, e.g., by suggesting fruitful questions and directions of inquiry), and to synthesize a new "best" alternative (or best compromise) from the options available. Some formal MCDP's perform several of these functions at once.

There are far too many formal MCDP's to describe them all here. They differ in their conceptualization and structuring of the problem, in the form and complexity of the mathematical procedures used, in the number of different alternatives (and/or decision criteria) that can be efficiently accommodated, in the inputs that must be supplied by the decisionmaker, and in the amount and type of information produced.

One useful way to classify the various formal MCDP's (especially from the standpoint of examining their potential value for police decisionmaking) is in terms of the role of the decisionmaker. There is, in fact, a spectrum of MCDP's ranging from purely formal procedures, which require virtually no subjective information from the decisionmaker, to iterative interactive procedures that are critically dependent upon decisionmaker inputs--in particular, the systematic, sequential elicitation of the decisionmaker's preferences. Between these two extremes lie the mixed procedures, which combine some subjective decisionmaker inputs with heavy reliance on purely formal decisionmaking techniques. In a "mixed" procedure, subjective information is elicited from the decisionmaker only once (rather than iteratively); this information then becomes the input for a mathematical formula or algorithm that proceeds to select the "best" alternative (or alternatives).

1. Purely Formal Procedures. These MCDP's "solve" the multi-criterion decision problem without requiring explicit inputs from the decisionmaker concerning his or her preferences. Once the performance criteria are specified, the available alternatives evaluated in terms of each criterion, and any related constraints, targets, or goals

are spelled out, the formal algorithm takes over to identify the "best" alternative(s). Such procedures are, in effect, <u>normative</u>: they incorporate an implicit standard of what constitutes the "best" alternative, a standard that, it is assumed, would be accepted by any rational decisionmaker. These standards are build into the mathematics of the procedure, and a decisionmaker must rely on the formal properties of the mathematical machinery (and the underlying assumptions and axioms--e.g. of Pareto optimality) to ensure that an appropriate ranking of alternatives is produced.

The following examples illustrate the kinds of purely formal procedures that have been devised for addressing multi-criterion decision problems:

- <u>Vector Maximum Programming</u>:[10] This technique involves a very general formulation of the multi-criterion decision problem--e.g., as the maximization of a vector-valued criterion function over a convex set of alternatives, subject to a set of constraints. The criteria and the constraints usually must be continuous functions, although they can be linear or non-linear. Solution procedures generally involve the application of Kuhn-Tucker conditions and similar general requirements to identify the <u>set</u> of non-dominated points or regions that constitute the "best" (e.g., efficient or Pareto optimal) alternatives.[11]
- <u>Multi-Objective Linear Programming</u>: This family of procedures represents a restriction of the vector maximum problem to the case of linear objective (criterion) functions and linear constraints on the values of the performance criteria. Most of these approaches utilize variations of the simplex algorithm to determine the optimal alternative or alternatives, e.g., by identifying all non-dominated extreme points in the convex set of alternatives defined by the problem. Examples of procedures that can be used when the alternatives are continuous include the "Revised Simplex Method" and the "Multicriteria Simplex Method."[12]

- Multiparametric Decomposition:[13] This procedure represents another way to approach the case of linear objective (criterion) functions with linear constraints. However, the focus here is on combining the various objective functions into one criterion, rather than keeping them separate (as is done in multi-objective linear programming). The k objective functions $f_i(x)$, where x is a vector characterizing an alternative, are combined by forming a weighted sum, $\sum \lambda_i f_i(x)$, where the λ_i correspond to a set of k non-negative criterion weights that sum to 1. It can be shown that if the foregoing weighted sum is maximized with respect to x over the convex space of feasible alternatives defined by the (linear) constraints, using all possible feasible combinations of criterion weights, one will generate the complete set of non-dominated extreme points, $\{x^o\}$, just as if one had used the multi-objective linear programming procedures described previously. But more importantly, with each non-dominated extreme point, say x^o_j, the multi-parametric decomposition procedure associates a subset of combinations of criterion weights, each of which will make that non-dominated extreme point optimal (e.g., by maximizing $\sum_i \lambda_i f_i(x)$ at $x = x^o_j$). This decomposition of the set of all possible weightings allows one to determine just which combinations (and ranges) of criterion weights are needed to make any given non-dominated solution "the best" - an important consideration for sensitivity studies and other analyses of the multi-criterion problem. Moreover, no *a priori* information on the relative weightings of the various objectives is needed to effect this decomposition of the criterion weights.
- Canonical Analysis:[14] This procedure relies on canonical correlation techniques to synthesize a single linear performance function that incorporates all the criteria. To use this approach, one must identify a set of performance criteria and a set of "predictor" variables that (it is assumed)

jointly determine the values of the performance criteria for any given alternative. If the actual values of the performance criteria and predictor variables are provided for each decision alternative, a canonical correlation can be performed to identify the linear combination of criterion variables and the linear combination of predictor variables that exhibit the highest mutual correlation for the given set of alternatives. The resulting linear combination of criteria can be used as a performance index for ranking the alternatives.

- Unit Weighting.[15] To utilize this procedure, the values of the various performance criteria for a given set of discrete alternatives must be "standardized"--that is, transformed so as to have zero mean and unit variance. For each performance criterion and alternative, the mean value of the criterion is substracted from the actual value of the criterion for the given alternative, and the result is divided by the standard deviation of the criterion. (The mean and standard deviation for a given criterion are computed from the criterion values for the entire set of alternatives.) THe standardized criteria for a given alternative are then added together, and the resulting index is used to rank the alternatives.

Note that the first three procedures described above all involve variations of the vector maximum approach, while the last two techniques emphasize the use of statistical methods for solving (or simplifying) the multi-criterion decision problem.

In addition, many of the "mixed" procedures described later can be reduced to purely formal MCDP's if the need for decisionmaker inputs can be avoided (e.g., by making appropriate assumptions or by using formal [mathematical] procedures to generate the needed inputs). For example, if one assumes that all deviations from performance targets or from ideal performance levels are equally important, no subjective decisionmaker inputs are needed to apply MCDP's involving goal programming or compromise programming (see below), and the latter techniques in effect reduce to purely mathematical exercises for identifying a best alternative or a compromise set of non-dominated alternatives from the given performance data. Or instead, one might use criterion weights derived from one of the many mathematical formulas that have been suggested for that purpose. For instance, Zeleny suggests using criterion weights based on the entropy--e.g., the decision information--contained in and transmitted by each criterion.[16] (The greater the variation among the alternatives with regard to a given criterion, the greater the entropy and hence the weight given to that criterion for purposes of distinguishing between alternatives. Conversely, if all alternatives exhibit virtually the same level with regard to a given criterion, the latter is assumed to be of little value in distinguishing between alternatives and is weighted accordingly.) Similarly, if quantitative performance data are available on each criterion for a set of alternatives, many of the ratios and matrices needed for the analytic hierarchy process (normally a "mixed" approach--see below) can be computed without the need for decisionmaker inputs, and the procedure reduces to a purely mathematical exercise.[17]

Of course, the reduction of a mixed MCDP to a purely formal approach involves a number of implicit assumptions concerning decisionmaker weights and priorities. Whether or not such assumptions are tenable--or, in fact, alter the ultimate outcome--depends on the specific case in question.

2. Iterative Interactive Procedures. Instead of relying on a set of implicit normative assumptions to mathematically determine the "best" alternative, iterative interactive techniques focus on identifying the decisionmaker's own implicit preferences for the alternatives in question and using those preferences to find the decisionmaker's most preferred option. In the typical iterative interactive procedure, the decisionmaker is asked relatively simple questions about his or her preferences concerning certain criteria or options, and the responses are entered into a

mathematical algorithm. The latter may organize the information, identify the major remaining information needs, suggest promising directions for further inquiry, and/or define the next round of questions for the decisionmaker. The result is that over time, the decisionmaker's responses are iteratively refined, supplemented, and synthesized until there is enough information on the decisionmaker's preferences to rank the various options. Alternatively, the decisionmaker's responses may be used to guide the algorithm itself in efficiently searching for a solution, e.g., by suggesting promising directions to explore or by limiting the number of alternatives considered.

Note that in an iterative, interactive multi-criterion procedure, the mathematical algorithm serves only to guide and facilitate the efficient elicitation, exploration, and use of the decisionmaker's preferences. Thus, the algorithm's own implicit priorities and assumptions have much less of a role in arriving at a solution of the multi-criterion decision problem than they do for purely formal MCDP's.

While there are a great many iterative, interactive procedures for solving the multi-criterion decision problem, most of the major ones are based on a relatively limited set of approaches. Thus, many such procedures focus on systematically eliciting and refining estimates of the decisionmakers's substitution rates--e.g., the amount of one criterion that the decisionmaker would be willing to trade for one more unit of another criterion, everything else being equal. This information is used to sequentially estimate criterion weights, to define the general "direction" of the decisionmaker's preferences, and/or to characterize the topology of the decisionmaker's utility function in the vicinity of certain alternatives--for instance, to direct a search algorithm seeking the point of maximum utility or to determine the relative position of specific multi-criterion alternatives on the decisionmaker's utility surface.

Other iterative interactive techniques emphasize direct, sequential reduction of the set of non-dominated solutions (the latter are often too numerous to be of much immediate help to a decisionmaker seeking his optimal point among them). These procedures focus on sequentially "pruning" the set of non-dominated solutions (with the decisionmaker's help) until the most preferred point is found.

Another group of iterative, interactive approaches makes use of the so-called "ideal" point - a (usually) hypothetical alternative for which each criterion achieves its "best" attainable value over the entire set of feasible alternatives. The ideal point can be used to provide the decisionmaker with a (sometimes evolving) basis for creating, assessing, and refining compromise solutions.

Other techniques - as well as various combinations of the foregoing methods - are also possible. Indeed, several of the purely formal approaches described previously, as well as the mixed techniques to be described later, can be adapted to an iterative, interactive format. The following examples are illustrative of the iterative, interactive approaches available for solving multi-criterion decision problems.[18]

- <u>Interactive Programming</u>:[19] This procedure searches the space of feasible alternatives, X, for the point (alternative) that maximizes the decisionmaker's overall utility function--e.g., the most preferred alternative. The overall utility (preference) function U is assumed to be unknown but differentiable, with positive marginal utilities. The various individual objective functions (performance criteria) $f_i(x)$ are assumed to be well-defined, known functions of the vector, x, which characterizes and identifies a given alternative. The decisionmaker's utility function is approximated as a linear expression:

  $$U(x) = w_1 f_1(x) + w_2 f_2(x) + \ldots + w_r f_r(x)$$

  near any given point x in the feasible space of alternatives X. With the help of the decisionmaker, the iterative programming algorithm determines the best direction and distance for moving from alternative x^k (currently being considered) to a new point x^{k+1}. The move is selected so as to provide the greatest increase in the (implicit) utility function. Assuming the above linear approximation to the decisionmaker's utility function, the best direction for increasing U can be computed from the gradient of U, approximated as $\sum_{i=1}^{r} w_i \nabla f_i(x)$ where $\nabla f_i(x)$ is the gradient of f_i at x,

$$\nabla f_i(x) = \left(\frac{\partial f_i}{\partial x_1}, \ldots\ldots, \frac{\partial f_i}{\partial x_r} \right)$$

The appropriate direction is determined and utilized through the following algorithm:

(1) Choose a starting point, x^1, within the feasible space of alternatives.

(2) At a given point, x^k (x^1 if this is the first iteration), estimate the gradient of U by first assessing the weights $w^k{}_i$ using information provided by the decisionmaker. These weights are the marginal substitution rates (e.g., indifference tradeoffs) between each <u>objective</u>, f_i, and an arbitrary reference objective, say f_1. To estimate these weights, the decisionmaker is asked how much he would be willing to reduce the value of the i^{th} objective, $f_i(x^k)$, to compensate for a small increase in the reference objective, $\Delta f_1(x^k)$. From the decisionmaker's response (e.g., the corresponding change $\Delta f_i(x^k)$ in the i^{th} objective), one can estimate the substitution rate as:

$$w^k{}_i = \frac{\Delta f_i(x^k)}{\Delta f_1(x^k)}$$

The other substitution rates can be estimated in like fashion.

(3) Next, an approximation for the directional derivative,

$$\sum_{i=1}^{r} w_i^k \nabla f_i(x^k) \cdot y^k$$

is maximized over all possible directions y^k within the feasible region, X. This represents a linear or non-linear programming problem, depending on the f_i. The maximizing vector, y^k, determines the best direction of movement, d^k (computed as $y^k - x^k$).

(4) The decisionmaker is then asked to select the optimal distance to move in direction d^k by examining various stopping points, x^{k+1}:

$x^{k+1} = x^k + t^k d^k$ (t^k is a scalar)

This can be done by evaluating each objective f_i $(x^k + t^k d^k)$ for various values of t^k until a satisfactory stopping point is found.

One then repeats steps 2-4 for point x^{k+1}, terminating when $x^{k+1} = x^k$ (or is close enough to x^k for practical purposes).

Dyer has proposed an analogous iterative interactive approach for solving goal programming formulations of the multi-criterion decision problem (goal programming is discussed below under "mixed" MCDP's).[20] The weights associated with deviations from the targeted criterion levels (the goals in the goal programming problem) can be estimated and iterated by the decisionmaker using steps 2-4 above.

- Zionts-Wallenius Method: Wallenius has found that decisionmakers experience considerable difficulty in estimating marginal rates of substitution between objectives (something that must be done to use the iterative programming procedures described above). The Ziont-Wallenius method is an attempt to address this problem. To simplify the interaction with the decisionmaker, the procedure assumes that the decisionmaker's implicit utility function, U, is a linear weighted sum of the various individual criteria functions, $f_i(x)$:

$$U(x) = f(\lambda, x) = \sum_{i=1}^{r} \lambda_i f_i(x)$$

where the λ_i are non-negative weights that sum to one. The solution procedure consists of systematically assessing whether it would be desirable to move from one non-dominated extreme solution to an adjacent non-dominated solution. If the answer is yes, the move is made and the process is repeated. The search is guided by eliciting the decisionmaker's assessment of the desirability of the various tradeoffs involved in moving away from a given non-dominated extreme solution. The original (and simplest) version of the Zionts-Wallenius procedure was generally as follows:[21]

(1) Select an arbitrary set of non-negative weights, λ_i, normalized to sum to one.

(2) Using this set of weights, maximize $\Sigma_i \lambda_i f_i(x)$ over the feasible set of alternatives, X (e.g., using linear programming procedures). The result will be a non-dominated extreme point (alternative), x^o.

(3) Next identify (from the simplex tableau) <u>all</u> non-dominated solutions adjacent to x^o.

(4) For each adjacent non-dominated solution point, determine (perhaps using the simplex tableau) the tradeoffs that would be involved in moving to that point from x^o (e.g., the increases and decreases that would occur in the various objective functions in connection with such a move). These tradeoffs (characterized as changes in performance) are then presented to the decisionmaker, who must assess the overall desirability of each move as yes (desirable), no, or indifferent/no judgment possible.

(5) If all responses are "no," the procedure terminates and the corresponding set of weights (and the non-dominated solution, x^o) are optimal. Otherwise, the "yes" answers are translated into new constraints on the

weights, and a new set of weights, consistent with the constraints, is selected.

(6) The problem given in Step 2 is then solved again, using the new weights. This yields a new non-dominated extreme solution, x^1.

(7) The decisionmaker is then asked to choose between x^o and x^1. If he chooses x^o, the latter is optimal (at least locally). If he chooses x^1, a new constraint is generated based on the decisionmaker's preference for x^1 over x^o, and one returns to Step 3 and repeats the procedure, this time using x^1 in place of x^o.

The foregoing procedure has been extensively refined since it was first proposed in 1975. The current version actually involves 13 steps and includes decisionmaker assessments of the relative overall desirability of adjacent extreme non-dominated solutions (e.g., pairwise point comparisons), in addition to the (previously described) assessments of the tradeoffs involved in moving to adjacent non-dominated solutions.[22] The refinements to the algorithm make it possible to handle more general classes of utility functions (they need not be strictly linear). They also respond to findings by Zionts and Wallenius in using the original procedure indicating that managers seem to prefer choosing between alternatives and evaluating tradeoffs.

Nevertheless, the procedure is still somewhat complex. A further disadvantage is that it requires "holistic" pairwise comparisons between multi-dimensional alternatives; however, the difficulties and unreliability of such (subjective) comparisons are precisely what led us to look to the various MCDP's for help.

- Stankard/Maier-Rothe/Gupta Procedure.[23] Unlike the foregoing techniques, this procedure does not require that the multi-criterion decision problem itself be formulated as a linear or non-linear programming problem. The algorithm deals directly with the alternatives of interest and their (given) performance attributes. On the other hand, a disadvantage of this approach is that it can deal with only two alternatives at a time.

The procedure operates by successively tightening estimates of the decisionmaker's substitution rates (which are known to lie within certain ranges) until one can show that for any possible combination of substitution rates within these ranges, one of the alternatives will always have a higher utility than the other (and hence is the preferred alternative).

The rationale for the procedure is as follows. For the two alternatives of interest (A and B), characterized by performance vectors x^A and x^B, consider the alternative that is midway "between" them in the sense that it is characterized by a performance vector $x^* = 1/2(x^A + x^B)$. It is assumed that the decisionmaker's preferences for these alternatives are reflected in an unknown (implicit) utility function, U(x). If (say) A is preferred to B, it can be shown that

$$\sum_{i=1}^{r} (x_i^A - x_i^B)\, s_i = Z(s_1, \ldots s_r) \geq 0 \qquad (1)$$

where s_i is the finite difference approximation to the marginal substitution rate between criterion x_i and an (arbitrary) standard criterion, say x_r, evaluated at the point x*:

$$s_i = \left. \frac{\Delta u / \Delta x_i}{\Delta u / \Delta x_r} \right|_{x^*}$$

The s_i must be non-negative (for positive increments, Δ). Each allowed combination of r substitution rates $(s_1, s_2, \ldots, s_r)$ represents a vector s, and the set of possible substitution rate vectors defines a domain D.

One can determine whether A is preferred to B as soon as one has enough information to determine the sign of the left-hand side of (1) for <u>all</u> feasible s_i, e.g., all s_i in the domain D. The objective of the algorithm is to successively restrict the size of D (using decisionmaker inputs) until the maximum and the minimum values of Z over D have the same sign--in other words, for all feasible substitution rates Z is positive (or negative). The algorithm proceeds as follows:

(1) Select one of the performance criteria, x_i, and ask the decisionmaker which of the following two alternatives he prefers:

$$\begin{bmatrix} x^*_1 \\ \cdot \\ \cdot \\ x^*_i + \Delta x_i \\ \cdot \\ \cdot \\ x^*_r \end{bmatrix} \text{ or } \begin{bmatrix} x^*_1 \\ \cdot \\ \cdot \\ x^*_i \\ \cdot \\ \cdot \\ x^*_r + \Delta x_r \end{bmatrix}.$$

where Δx_i and Δx_r are specified increments.

(2) On the basis of the decisionmaker's response, one can specify certain bounds on the s_i. Thus, if the left-hand vector is preferred,

$$s_i \geq \Delta x_r / \Delta x_i$$

If the right-hand vector is preferred, then

$$s_i \leq \Delta x_r / \Delta x_i$$

(3) The inequalities identified in Step 2 are then added to the definition of the domain D, and the minimum and maximum values of Z are again computed over that (now more restricted) domain. (These computations involve two rather simple linear programs that can usually be solved by inspection.)

(4) If the minimum and the maximum values of Z are both positive, then alternative A is preferred to alternative B. If the minimum and maximum values are both negative, then alternative B is preferred to A. If the signs differ, the bounds on the substitution rates are not yet tight enough to establish the direction of the decisionmaker's preference, and one must ask the decisionmaker to give his preferences for additional tradeoffs of the type shown in Step 1, using other performance criteria x_i and/or different increments Δx_i. The algorithm itself selects the "best" tradeoff question(s) to pose to the decisionmaker in the next round of questioning.

With these new choices of x_i and Δx_i, one returns to Step 1. The process continues until the domain D has been constrained enough to produce a clear determination of the decisionmaker's preference for A vs. B.

- Steeer's Interactive Procedure:[24] This rather complex approach utilizes interval programming along with progressive reduction of the "gradient cone" formed from the gradients of the various objectives. The decision problem is formulated as a multi-objective linear programming problem in which one is attempting to maximize a weighted linear combination of the objective functions (as discussed previously under multi-parametric decomposition). However, instead of maximizing the weighted combination of objectives over all possible sets of normalized weights, in this case the decisionmaker first specifies (to the extent possible) upper and lower bounds on each criterion weight. (A precise estimate of the criterion weights does not need to be elicited.) If the multi-criterion decision problem is now solved with each criterion weight restricted to the intervals initially prescribed by the decisionmaker, fewer non-dominated extreme solutions will result.

(The solution procedure is, however, rather complex; multi-objective linear programming in its usual form cannot be used.) The decisionmaker is presented with a sample of the resulting non-dominated extreme points and selects the most preferred point. The algorithm then alters the bounds on the criterion weights so that the most preferred point can be generated using a smaller subspace of weights. The result is another sample of non-dominated extreme points, but this time clustered around the decisionmaker's most preferred point. The decisionmaker now selects a most preferred extreme point from this new set; the bounds on the criterion weights are contracted once again; and the efficient points in an even smaller neighborhood of the revised "most preferred point" are generated. In the words of Steuer, "by the selection of one of these efficient extreme points, the algorithm is informed as to what part of the efficient surface should be explored in more detail. In this way, the algorithm iteratively 'focuses-in' with greater powers of magnification on the corner points of the feasible region surrounding the efficient extreme point of greatest utility."[25]

- The Displaced Ideal and Compromise Programming.[26] These two procedures (which are closely related) are conceptually quite different from the techniques presented previously. In principal, they are supposed to proceed iteratively (hence, their inclusion here). However, in may decision situations likely to arise in connection with multi-criterion assessments of overall police performance, these procedures will require only a single iteration; in such cases, they can be considered examples of the "mixed" approaches described later.

Both procedures attempt to narrow the set of non-dominated efficient solutions (alternatives) by measuring the "distance" of each alternative from a (hypothetical) "ideal" point and focusing only on those alternatives that are, in some sense, "closest" to the ideal. The "ideal" alternative is defined as a composite of the highest (best) performance scores attainable for each attribute. The ideal can sometimes be specified in absolute terms by consideration of each particular criterion--e.g., zero complaints, 100 percent of the survey respondents rating police performance as "excellent," etc. More often the ideal is defined in <u>relative</u> terms as the maximum of each objective (performance) function $f_i(x)$ <u>individually</u> attainable over the feasible (e.g., available) set of alternatives, X:

$$f^* = \text{"Ideal" Point} = (f^*_1, \ldots, f^*_r)$$

$$f^*_i = \underset{X}{\text{Max}}\ f_i(x), \quad i = 1, \ldots, r$$

Clearly, as the feasible set of alternatives X changes, the specifications of the ideal point may also be altered.

The "distance" measure which is used is of the form

$$d_p = \left[\sum_{i=1}^{r} \lambda_i^{\,p} (y^*_i - y_i)^p \right]^{1/p} \qquad 1 \leq p \leq \infty$$

where y^*_i is the ideal value of criterion i, y_i is the actual value of the criterion for a given alternative, the λ_i are a set of non-negative weights that sum to one, and p is a parameter that determines just how distance is being measured. For $p = 1$, the measure gives the longest distance in a geometric sense--e.g., all deviations are added, as though one were constrained to walk along city blocks in measuring distance. For $p = 2$, one has the usual Euclidean distance measure, the shortest distance between two points as measured along a straight line. For $p = \infty$, distance is measured as the maximum weighted deviation of <u>any</u> objective from its ideal value:

$$d_{\infty} = \text{Max}\ \lambda_i\ (y^*_i - y_i),\ i = 1, 2, \ldots, r$$

Note that maximization of the above distance measure leads to the familiar "min-max" expression: $\underset{X\ \ i}{\text{Min Max}}\left\{\lambda_i\ [y^*_i - y_i(x)]\right\}$.

To avoid scaling problems, the deviations from the ideal in (2) can be renormalized so that all deviations are commensurate. The resulting distance measure is:

$$L_p(x) = \left(\sum_{i=1}^{r} \lambda_i^p \left[\frac{f_i^* - f_i(x)}{f_i^*}\right]^p\right)^{1/p} \quad p = 1, 2, \ldots, \infty \qquad (3)$$

The usual procedure followed in connection with the displaced ideal and compromise programming approaches is to compute the distance L_p for p = 1, 2, and ∞ for each non-dominated point (or for a sample of such points). In general, there will be a different non-dominated point minimizing L for each p. These points are termed "compromise" solutions--e.g., the closest feasible alternatives to the ideal (in the sense given by p--e.g., a "city block" sense, a "straight line" sense, or a "min-max" sense). The decisionmaker may decide to accept one of these points as optimal. Alternatively, he may decide to delete some of the most "distant" points. If the "ideal" has been defined in a relative sense, elimination of certain alternatives may alter the maximum criterion values achievable over the remaining alternatives and hence change (displace) the definition of the ideal point. If this happens, the distances between the remaining alternatives and the displaced ideal are computed once again for the three values of p, and the process is repeated. The procedure ends when the pruning of "distant" alternatives by the decisionmaker no longer displaces the ideal, and the various distance measures stabilize. At that point, the decisionmaker makes a final selection of his most preferred "compromise" solution from the alternatives "closest" to the ideal according to L_1, L_2, and L_∞.

Note that the decisionmaker must initially supply weights λ_i indicating the relative importance of deviations from the ideal for each objective. Or alternatively--as noted previously--various mathematical formulas (e.g., based on entropy considerations) can be used to relieve the decisionmaker of the responsibility of coming up with an a priori weighting.

Several other iterative, interactive procedures rely upon the identification of an "ideal" point and/or the distance to such a point for guiding the search for a best compromise. Of particular interest are the so-called Progressive Orientation Procedure (POP) and its descendent, STEM.[27]

3. Mixed Procedures. These techniques represent a compromise between the purely formal MCDP's (which incorporate virtually no information on individual decisionmaker preferences) and the iterative interactive methods (which rely heavily on such information). In the case of "mixed" procedures, the decisionmaker is involved but only once: decisionmaker inputs are needed only to initiate the algorithm, after which the procedure mathematically selects the best alternative (or alternatives) without further interaction with the decisionmaker. Nevertheless, the effort required of the decisionmaker in connection with mixed procedures can be quite substantial, depending on the specific mixed MCDP used and the nature of the decision problem (e.g., the number of alternatives and/or criteria involved). The following examples illustrate the major types of "mixed" procedures currently in use.

- Goal Programming:[28] This procedure employs linear programming techniques to identify (or synthesize) an alternative that comes "as close as possible" to a set of simultaneously incompatible goals. A desired level (the goal) is specified for each criterion in advance, along with a vector-valued linear function that characterizes the performance of each alternative in terms of the various criteria. The decisionmaker's role consists of initially specifying the weights for under- and over-achievement of each goal. These weights can be analogous to the weights used in connection with the weighted sums of performance criteria discussed previously, or they can take the form of "pre-emptive" (priority) weights which, in effect, define a hierarchy among the various criteria: the highest priority goals are satisfied first, and only then are lower-priority goals addressed (but always without disturbing the attainment of the higher priority goals). Regardless of the type of weights provided by the decisionmaker, the goal programming problem can be formulated as a linear programming exercise, and appropriate linear programming routines can be used to identify the alternative that (for example) under-achieves the specified goals to the smallest degree.

- Interval Programming:[29] This procedure was previously mentioned in connection with Steuer's interactive programming approach, where it was used in an interactive format. Interval programming can also be used as a mixed (one-shot) technique for reducing the number of non-dominated solutions associated with a given multicriterion decision problem. To do this, the decisionmaker is initially asked to specify a range (interval) for the weight (e.g., the importance) of each criterion or objective function. It should, in fact, be much easier for the decisionmaker to provide interval estimates of the criterion weights than point estimates or even tradeoffs (substitution rates). The tighter the estimates (and the intervals within which the weights are believed to lie), the fewer the non-dominated solutions that are generated. The solution procedure is, however, rather complex. It involves using the gradient cone (formed from the gradients of the various objectives) to convert the interval programming problem into a vector maximum problem, for which computerized solution procedures are available.

- Point Allocation:[30] This is a very simple technique for developing an index to rate the performance of each alternative, although it lacks a formal theory. The decisionmaker is merely told to allocate 100 points among the k performance criteria used to evaluate the various decision alternatives. The 100 points are to be allocated so as to reflect the relative importance of each criterion from the decisionmaker's perspective (e.g., if criterion 1 is given 20 points and criterion 2 receives 10 points, criterion 1 is twice as important as criterion 2). The performance criteria are then scaled to be commensurate (e.g., by dividing each criterion by its maximum value over all available alternatives). A performance index is computed for each alternative by multiplying each scaled performance score by the points (importance) assigned to it and adding the results. The alternative with the largest index is assumed to be "best."

- Keeney-Raiffa Method:[31] This technique generates an estimate of the decisionmaker's entire utility function, U, as a function of the various performance criteria, x_i. To evaluate the relative worth of any given alternative, one merely substitutes the performance characteristics for that alternative into the utility function and computes the associated utility. The latter figure can be used to rank or compare the given alternative with other alternatives, and the alternative with the highest computed utility is assumed to be the "best."

The Keeney-Raiffa technique has been widely described and applied; indeed it has often been considered THE multicriterion decision procedure. The technique is also distinguished by an extensive theoretical foundation.

To execute the Keeney-Raiffa technique, one must make some strong assumptions concerning the structure and properties of the decisionmaker's preferences. The validity of these assumptions must be verified with the decisionmaker before proceeding with the technique.

The assumptions allow one to decompose the utility function into separate, easily-estimated parts. (This is analogous to the probability independence assumptions used to simplify probabilistic models.) The following are illustrative of the kinds of assumptions that are usually made and their consequences:[32]

> <u>Assumption 1</u>: If all alternatives have identical performance scores on every criterion but the i^{th}, then the decisionmaker's preferences concerning differences between any two alternatives with regard to the i^{th} criterion will be unaffected by changes in the levels of the other performance criteria.

In other words, a sort of "preference independence" holds: the level of one criterion (say, crime rates) does not affect a decisionmakers's preferences for alternate levels of <u>other</u> criteria (for instance, quality of arrest), everything else being equal. The <u>only</u> type of utility function consistent with Assumption 1 is an additive, separable function of the form

$$U(x) = \sum_{i=1}^{r} w_i p_i(x)$$

where the w_i are non-negative weights that add to 1, and the p_i's are unidimensional preference functions for the x_i (the p_i are scaled so that their values lie between 0 and 1).

> <u>Assumption 2</u>: If all alternatives have identical performance scores for a subset of criteria, T, then decisionmaker preference orderings for differences in performance between the alternatives depend only on the

performance scores for the (complementary) subset I of the remaining criteria.

If Assumption 2 holds, then the decisionmaker's utility function must be given by either:

$$1+\lambda u(x) = \prod_{i=1}^{r} [1+\lambda p_i(x_i)] \quad \text{if } \lambda \neq 0 \qquad (4)$$

or

$$u(x) = \sum_{1}^{r} w_i p_i(x_i) \quad \text{if } \lambda = 0 \qquad (5)$$

where again the $p_i(x_i)$ are unidimensional preference functions that vary between 0 and 1, and λ is a scaling constant greater than -1 and defined by the equation:

$$1 + \ = \prod_{i=1}^{r} (1 + \lambda w_i) \qquad (6)$$

Assumption 2 implies that the utility function is either additive or multiplcative. The multiplicative form allows for interdependencies among the criteria through the constant, λ.

To utilize the Keeney-Raiffa approach, the following steps must be completed:

(1) Verify that the appropriate assumption (e.g., 1 or 2) holds for the decisionmaker. Several authors have suggested questions and dialogues that can be used to establish the validity of such assumptions.[33]

(2) Assuming that Assumption 1 or 2 holds, specify the univariate preference functions $p_i(x_i)$ for each criterion x_i. A number of procedures can be used to accomplish this—e.g., by scaling existing performance criteria so that their values lie between 0 and 1, by assuming a general form for the p_i and calibrating it from the decisionmaker's responses to certain questions, etc.

(3) Evalute the r weights w_i. To do this, one develops r equations in the r unknowns, w_i, and solves them simultaneously for the w_i. (If

Assumption 2 applies, these equations may be non-linear and their solution quite difficult.) Keeney and Raiffa propose that the necessary equations be specified by asking the decisionmaker for his preferences concerning certain lotteries over specific values of the performance criteria. Using the $p_i(x_i)$ estimated in Step 2, the necessary equations can then be synthesized and solved for the w_i.

(4) If the multiplicative utility function is being used, estimate the tradeoff parameter λ by evaluating the utility function at the ideal point, x* (where all performance criteria are at their "best" or most preferred levels). Then

$$1 + \lambda = \prod_{i=1}^{r} (1 + \lambda w_i)$$

which can be solved (iteratively) for λ, given the w_i estimated in Step 3.

(5) At this point, the utility function U (x) will be completely specified, and one need only substitute the actual performance values x_i for a given alternative to find the corresponding utility of that alternative. The resulting utility figures for the various alternatives can be used to rank those altrnatives from most preferred to least preferred.

There have been a number of criticisms of the Keeney-Raiffa approach. In the first place, the procedure is very complex and time-consuming, putting great demands on the time and patience of the decisionmaker. Furthermore, a great deal of information is collected but then lost in "collapsing" to a single utility figure. This loss of information may obscure potentially important differences between the alternatives. In addition, the utility functions are difficult to update over time or when circumstances change. And finally, the intractibility of the procedure increases rapidly as the number of performance criteria increases; Starr and Zeleny claim that the procedure becomes impractical for more than 3-5 attributes.[34]

- <u>Multi-Attribute Utility Technology (MAUT)</u>:[35] This procedure can be viewed as a simplification of the various judgments and decisionmaker inputs required in connection with the Keeney-Raiffa Method. While the procedure requires much less effort than the Keeney-Raiffa approach and is relatively easy to understand, these simplifications are achieved at the cost of introducing several strong assumptions, e.g., on the linearity and separability of the utility function, among other things. MAUT involves the following steps:

  (1) Identify the "stakeholders," persons making or affected by the decision(s) of interest.

  (2) Organize the performance criteria into a hierarchy or "value tree" that reflects the relationships of the various criteria. The lowest level of the hierarchy will consist of the individual performance criteria. At the next level, all criteria that contribute to a common subgoal are grouped under that subgoal. Groupings of subgoals under common overarching goals constitute the next level of the hierarchy, and so forth until the tree terminates in a single overall objective.

  (3) Have the stakeholders assess the relative importance of each criterion (the lowest level of the hierarchy). To do this, weights are assigned to the branches at <u>each</u> node of the value hierarchy (the weights associated with a node must sum to one). These weights are synthesized subjectively, using pairwise comparisons among the criteria at a given node. Procedures are available for identifying inconsistencies in the weights and for combining weights from different stakeholders (e.g., by averaging, sensitivity analysis, or the use of non-mathematical group techniques such as the Delphi Method). To find the relative importance

of each individual performance criterion (at the lowest level of the hierarchy), one multiplies together the weights associated with each branch of the value tree in the path from the top of the hierarchy to the given criterion at the bottom.

(4) Assess the performance of each alternative with respect to each of the criteria and scale them for commensurability. This can be done using actual (objective) performance measures or subjective ratings (e.g., by experts). The performance measures (known as "location measures" in MAUT terminology) must be scaled to reflect the value or utility that the decisionmakers (or stakeholders) associate with specific performance levels. The rescaling is also designed to make the various performance measures (and their associated utility measures) commensurable (the rescaled measures all use a scale of 0 to 100). A variety of practical scaling procedures can be used, including linear and bilinear transformations, subjective assessments, etc. In some cases, the rescaling process has been facilitated by letting the stakeholders first select the <u>form</u> of the scaling transformation (from several standard shapes) and then calibrating the relevant curve at a few points.

(5) Compute the overall utility of each alternative by multiplying each scaled performance (location) measure by the weight associated with the corresponding criterion and adding the results. These utility figures are then used to rank the alternatives.

(6) Finally, a sensitivity analysis is conducted to assess the impacts of errors or changes in any of the weights, measures, or assessments involved.

- Analytical Hierarchy Process (AHP):[36] This procedure is somewhat different from the other mathematical approaches that have been described. It is based on pairwise comparisons of criteria and of alternatives to assess their relative importance or their contribution to the various performance criteria. The results of the pairwise comparisons are used to construct a number of special matrices, from which one can derive estimates of the criterion weights and the overall contribution of each alternative to each criterion. These data are then combined (as a linear weighted sum) to produce an overall priority index for each alternative.

Assume that there are r performance criteria, denoted as C_1, C_2, ... , C_r, and n alternatives denoted by A_1, A_2, ..., A_n. It is also assumed, to simplify the presentation, that all performance criteria contribute to a single overall objective, e.g., control of crime. Application of the analytic hierarchy process to this situation involves the following steps (for the theoretical rationale behind these steps, see the references given in footnote 36):

(1) Have the decisionmaker compare and rate each pair of performance criteria in terms of their relative importance with regard to the overall objective of controlling crime. The following reciprocal scale should be used:

Very much more important	9
Much more important	7
Moderately more important	5
Slightly more important	3
Equally important	1
Slightly less important	1/3
Moderately less important	1/5
Much less important	1/7
Very much less important	1/9

(2) Construct a matrix of these importance ratings. Note that if element a_{ij} of this matrix (the importance of C_i vs. C_j) has a value of x, then

element a_{ji} (the importance of C_j vs. C_i) has a value of 1/x. Denote this k x k matrix of pairwise comparisons by C.

(3) Find the largest eigenvalue of the C matrix, denoted by λ_{max}. In addition, find the eigenvector associated with λ_{max}, and normalize its components so that they sum to one. Denote this eigenvector by w, with elements w_i, i = 1, . . ., r. The w_i can be identified with the weights (the relative importance) of the various performance criteria (e.g., w_i is the weight of the i^{th} criterion).

(4) Next, compare the importance of each alternative with that of every other alternative as regards their contribution to criterion C_1. These pairwise comparisons can be conducted in either of two ways: (1) by having the decisionmaker make a subjective rating using the reciprocal scale given under Step 1, or (2) by computing the appropriate ratios using actual performance data (when such information is available). (For instance, if C_1 is the clearance rate for Part I crimes and alternative A_1 had a clearance rate of 12 percent while alternative A_2 had a clearance rate of 24 percent, the relative importance of A_1 vs. that of A_2 with regard to clearing Part I crimes would be computed as 12/24 = 1/2; the corresponding importance of A_2 vs. A_1 would be 24/12 = 2.) Note that if ratios of actual performance criteria are used, the ratios should be scaled so that none are greater than 9 (to make the scale analogous to the subjective scale used in Step 1; the scaling is also needed to improve the stability of the eigenvector computations in Step 6).

(5) Using the importance ratios or ratings obtained in Step 4, construct a matrix similar to that of Step 2. An element f^1_{ij} of this matrix corresponds to the relative importance of alternative A_i vs. A_j with

regard to criterion C_1. Once again, $f^1_{ji} = 1/f^1_{ij}$. Denote this n x n matrix of pairwise comparisons by F^1.

(6) Find the largest eigenvalue of matrix F^1, λ^1_{max}, and compute the corresponding eigenvector. Normalize the elements of the eigenvector so as to sum to one, and denote the results by g^1 (with elements g^1_1, . . ., g^1_n). Element g^1_i of the normalized eigenvector corresponds to the overall contribution of alternative A_i to criterion C_1.

(7) Repeat steps 4 - 6 for each of the remaining r-1 performance criteria--e.g., make pairwise comparisons of the various alternatives regarding the importance of their contribution to criterion C_2, their contribution to criterion C_3, etc. The result will be a series of pairwise comparison matrices F^2, F^3, . . ., F^r. For each matrix there will be a normalized eigenvector g^2, g^3, . . ., g^r.

(8) The r normalized eigenvectors g^i can be adjoined to form an n x r matrix G whose columns are the eigenvectors g^i. To determine the overall importance of each alternative, one multiplies matrix G by the normalized vector of criterion weights, w (from Step 3). The result is an n-dimensional vector V whose elements represent the values of the corresponding alternatives:

$$V = G\,w,$$

or

$$v_i = \sum_{j=1}^{r} w_j g^j_i$$

The v_i can be used to rank the alternatives with regard to their overall importance or priority. Note that, in the end, a simple weighted sum is used to compute the overall value of each alternative.

The analytic hierarchy process is quite tolerant of inconsistencies in the decisionmaker's assessments of the relative importance of the various alternatives and criteria. If the ratings have been consistent, then the following relationship holds for any triplet of matrix elements:

$$a_{ij} = a_{ik}a_{kj}, \text{ all } i, j, k$$

To the extent that the above relationship does not hold, the decisionmaker has been inconsistent in providing his ratings. The AHP can be used when such inconsistencies arise (which is common). Indeed, the procedure generates an index of consistency for a subjectively estimated matrix (the maximum eigenvalue minus n, the dimension of the matrix)/(n-1). This index is zero if the matrix is consistent.

- Social Judgment Theory (SJT):[37] This approach involves a statistical analysis of assessments that the decisionmaker has made when faced with other multicriterion decision problems. Through the use of regression analysis, the procedure attempts to identify and calibrate a statistical model of the decision outcome as a function of the various performance criteria. Once the model is specified, it can be used to assess new alternatives. The following basic procedure is used:
  (1) Develop a set of "test" alternatives that can be used to estimate and calibrate the model. Each alternative should exhibit realistic values for each performance criterion of interest, and the entire set of alternatives should provide a representative range of values for each criterion. Ideally, 30 to 40 such "test" alternatives should be developed.
  (2) The decisionmaker then rates the overall value or utility of each "test" alternative on the basis of the performance criteria provided.
  (3) A multiple regression analysis is conducted next, using the decisionmaker's ratings as the dependent variables. The regression equations

and regression coefficients that are developed relate the values of the performance criteria to the decisionmaker's overall ratings. A number of different functional forms are usually tried, e.g., the basic linear model, a logarithmic model, exponential models, conjunctive (product) models, etc.

(4) The regression results are then analyzed to determine the consistency of the ratings predicted by the model and the actual ratings given by the decisionmaker, the goodness of the regression fit, the face validity of the regression model, and whether the assumptions necessary to conduct the regression analysis are satisfied by the given data set (e.g., concerning the statistical independence of the various performance criteria).

Once a satisfactory model of the decisionmaker's preferences has been found, it can be used to assess new alternatives whose criteria lie within the ranges used to estimate the model.

- Compromise Programming: This procedure was described in some detail under iterative, interactive MCDP's. In effect, compromise programming identifies the non-dominated solution or solutions that are "closest" (in various senses of the word) to the "ideal"—a hypothetical alternative in which each performance criterion exhibits the highest or best level found in any of the available alternatives. (See the discussion under "iterative, interactive procedures" for the mathematical details of this approach.)

Note that if the decisionmaker chooses not to (or is unable to) iteratively alter (displace) the ideal point (e.g., by removing certain alternatives from consideration in view of their "distance" from the ideal), the compromise programming algorithm will not proceed iteratively. In that case, compromise programming reduces to the following "mixed" procedure:

(1) Define the ideal point relative to the available alternatives.

(2) Have the decisionmaker specify the weights λ_i indicating the relative importance of a deviation from the ideal with regard to the i^{th} criterion.

(3) Identify the efficient, non-dominated alternatives associated with the given multi-criterion problem. These can be found using vector maximum programming or--if there are a limited number of discrete alternatives--by inspection.

(4) Compute the generalized distance measure $L_p(x)$ (see equation 3) for each non-dominated alternative (or, if the set of alternatives is continuous, for a representative sample of non-dominated alternatives). The distance between each alternative and the ideal should be computed for $p = 1$ (e.g., with distance measured in the "city block" sense), $p = 2$ (with distance measured in the usual Euclidean sense), and $p = \infty$ (for a min-max distance measure).

(5) Select the alternative(s) that minimize L_1, L_2, and L_∞. In general, these will not all be the same alternative, but only a small number of alternatives will be involved. These are the "best" compromise alternatives in the sense of being "closest" to the ideal point. If desired, the decisionmaker can then choose an overall "best" compromise from among these alternatives.

Note that in this version of compromise programming, the decisionmaker provides inputs only at the beginning of the procedure (Step 2) and--possibly--in Step 5. The rest of the process is purely mechanical and governed solely by the mathematics. Thus, this version of the procedure qualifies as a "mixed" MCDP. Indeed, the "mixed" form of compromise programming seems especially suitable for multi-criterion decisions involving police performance--situations likely to involve only a few (discrete) alternatives, all or most of which are non-dominated (as can be determined by inspection).

Clearly, there are a great many formal (mathematical) approaches for solving the multi-criterion decision problem. In contrast to the unidimensional and the non-mathematical group techniques described previously, formal mathematical MCDP's appear to have a number of potential advantages in solving multi-criterion decision problems:

- The mathematical procedures are capable of encoding and making systematic use of a great deal of performance data, often in its original form (e.g., without having to apply arbitrary heuristics, ad hoc simplifications, or questionable transformations to prepare the data for analysis).
- These procedures usually make explicit the weights that the decisionmaker is assigning to the various criteria. This is important information in its own right (for sensitivity analyses as well as for understanding the decision).
- The procedures are explicit, testable (e.g., via sensitivity analyses), and --in most cases--rational, with substantial theoretical underpinnings.
- Mathematical MCDP's can usually handle as many as 8 to 10 different performance criteria, sometimes more (for instance, the number of criteria that can easily be handled by MAUT is virtually unlimited).
- For some techniques (e.g., social judgment theory and the Keeney-Raiffa approach), the procedure produces a stable formula that can be readily applied to similar situations in the future, without having to recompute the parameters or go back to the decisionmaker for additional inputs.
- For the mixed and iterative interactive approaches, participation in the decision process can be a valuable experience in its own right, e.g., by forcing the decisionmaker to carefully and systematically reflect on--and

clarify--criterion weights, assumptions, tradeoffs, and preferences that were previously recognized and dealt with only implicitly, if at all.

On the other hand, a number of potential disadvantages (at least from the standpoint of public administrators) have also become apparent in connection with some of the mathematical procedures used for dealing with the multi-criterion decision problem:

- Some of the techniques—e.g., point allocation, unit weighting—appear to have little underlying theoretical basis and give the impression of being rather ad hoc responses to the multi-criterion decision problem.
- A number of the techniques are quite complex, making it doubtful that public administrators would understand them, or at least would accept them as a valid and reliable guide for their decisions.
- The burdens placed on the decisionmaker by some of these procedures (especially the Keeney-Raiffa method) can be quite high in terms of the time required and the difficulty of the judgments called for.
- Some of these approaches--especially the purely mathematical techniques--seem to keep the decisionmaker at arm's length, allowing for little or no decisionmaker input that could better tailor the process (and the decision outcomes) to the decisionmaker's own preferences.
- Some of the procedures are so complex that a department would have to hire costly outside expertise to run them (although the development of microcomputer packages for some of the procedures--e.g., the analytic hierarchy process--may temper such criticism).

Given the great variety of mathematical procedures available, it is not surprising that each approach seems to reflect a different mix of advantages and disadvantages. To assess the most promising approaches for regular police department use, one must examine the specific mix of requirements associated with the application of a mathematical MCDP to the problem of regularly monitoring police department performance. Before doing so, however, it is instructive to review some of the recent applications of mathematical multi-criterion decision approaches to police and other public sector decision problems.

Applications of Mathematical Multi-Criterion Decision Procedures to Police and Other Public Sector Problems

Despite the potential disadvantages listed above, mathematical MCDP's have been applied to a number of public sector problems and agencies, including the police. The following is a sample of these applications.[38]

Police and Criminal Justice Applications. Bodily utilized the Keeney-Raiffa approach to combine several criteria (a measure of response time and a measure of workload equity)--and the preferences of different interest groups--in designing police patrol sectors.[39] MAUT, the simplified version of the Keeney-Raiffa technique, has been used in an evaluation of the Federal Community Anti-Crime Program[40] and in an assessment of the Office of Rentalsman as an alternative to the courts for tenant-landlord disputes.[41] Social judgment theory has been used to help the Denver police select handgun ammunition.[42] And Grizzle has used the analytic hierarchy process to examine the priorities that selected professional and lay groups place on various performance criteria used for assessing the performance of probation/parole agencies.[43]

Other Public Sector Applications. There have been a number of other public sector applications of the Keeney-Raiffa procedure, usually in connection with the evaluation of a specific program or option. Examples include assessments of nuclear power plant siting,[44] airport locations,[45] and air pollution policies.[46] (The Keeney-Raiffa procedure has also been proposed as an aid to setting private sector productivity goals.[47]) MAUT has been used to combine a variety of performance criteria in assessing the performance and productivity of the street maintenance department in Morgantown, West Virginia[48] and for ranking alternative desegregation plans for Los Angeles schools.[49] Public sector applications of the analytic hierarchy process include strategic transportation planning in the Sudan, health care planning, the setting of landuse priorities, and the assessment of educational policies.[50] Social judgment theory has been used to assess the overall organizational performance of a public organization[51] and to develop priorities for acquiring open space.[52] And compromise programming has been used in connection with regional planning in the Netherlands.[53]

While the above examples suggest that there have been a number of recent applications of formal multi-criterion decision procedures in the public sector, most of them have been "one-shot" efforts focusing on a specific narrow program or issue. Only three applications were related to the use of MCDP's in connection with the general problem of regular agency performance measurement (the application of MAUT to the assessment of street maintenance department performance in Morgantown, West Virginia; the use of the analytic hierarchy process for the examination of interest group priorities in connection with performance measures for probation/parole agencies, and the utilization of social judgment theory in developing a consolidated measure of the overall performance of a public agency). But overall, there is little evidence that mathematical MCDP's are being utilized by government decisionmakers on a regular basis. Indeed, most of the applications cited above were executed by consultants and academics--and they seemed to be of mostly academic interest. Mathematical MCDP's have not yet become a tool widely accepted and utilized by public administrators.

Choosing Mathematical Mutli-Criterion Decision Procedures for Use in Connection with Regular Performance Measurement

If it is to be used by government administrators in connection with the regular monitoring and assessment of overall agency performance, a mathematical MCDP will have to satisfy a number of requirements. Recognition of these requirements can help narrow the range of procedures that need to be considered. Two kinds of conditions are of interest: formal (mathematical) considerations, and user considerations.

Formal (Mathematical) Considerations. The nature of the decision problems typically faced in connection with efforts to regularly monitor agency performance imposes some important constraints on the mathematical procedures to be used.

- The decision alternatives to be considered will usually be discrete (rather than continuous) and limited in number (e.g., agency performance for several previous years, or perhaps the results of several specific programs). The number of points will probably be small enough that a sophisticated search effort will not be needed since complete enumeration should be possible. (It should also be possible to identify non-dominated solutions by inspection.)
- Performance data and results will only be available as ex-post point estimates. This implies that (1) uncertainty--and the ability to handle it--will not be much of a consideration in selecting an MCDP (the primary source of uncertainty will be measurement error), and (2) little will be known about the functional relationships between performance and other characteristics (no nice, differentiable formulas, objective functions, or other relationships are likely to be available).
- While the number of alternatives is likely to be relatively small, the number of criteria is expected to be relatively large--e.g., 8, 10, or more (especially in police departments, where, as noted in Exhibit 1, a great many good performance criteria are usually available on a regular basis). The criteria, moreover, are likely to be very diverse--survey results, recorded statistics, subjective assessments, qualitative ratings, etc. The MCDP must be able to handle them all.

- The performance criteria will often be strongly interactive and/or at least partly redundant. Inconsistencies in preference assessments must also be expected and addressed.
- Police (and other agency) performance criteria will not always exhibit a positive marginal utility (e.g., the conviction that more is better) over their entire range. There will, instead, be a saturation point at which the desirability of a given performance characteristic levels off or even declines with further increases in the level of performance.
- Performance assessments are likely to be repeated periodically--monthly, quarterly, or annually. This means that administrators will be periodically facing the same types of problems and decisions--how to reconcile and combine a certain set of performance criteria into an overall assessment of agency performance, whether this month's performance was better than last month's, etc. Thus, in principle it should be possible to re-apply a multi-criterion decision procedure with little additional effort, once it has been developed and calibrated. In part, this means that the procedures should not depend upon the specific set of alternatives being considered.
- Finally, the procedure should be able to handle multiple perspectives without blurring their differences in ways that mask the richness of that diversity.

User Considerations. In addition to the formal implications of using multi-criterion decision procedures in connection with regular monitoring of agency performance, there are a number of other--practical--aspects that must be considered from the standpoint of the user:

- Validity and accuracy: to what extent does the procedure accurately reflect the decisionmaker's preferences? Obviously, this is of central importance in connection with the acceptance of the procedure by the decisionmaker.

- Ability to adequately discriminate between alternatives: to what extent is the procedure able to reduce the possibilities to at most one or a very few "best" alternatives? The types of alternatives likely to be present in assessing police department performance are, for instance, likely to consist of mostly non-dominated solutions (given the large number of performance criteria involved). Thus, a procedure like vector maximum programming is unlikely to be of much help in reducing the number of promising options.
- Feasibility of conducting a sensitivity analysis: given the many assumptions needed and the critical role often played by subjective assessments, it is important to be able to test the sensitivity of any solutions with respect to changes in the various parameters. Procedures where sensitivity analysis is difficult or impossible (e.g., some of the group decisionmaking techniques) will be at a disadvantage.
- Burden on the decisionmaker: MCDP's that impose heavy burdens on the decisionmaker are unlikely to find wide acceptance. Several types of burdens are important. One is the volume of input needed from the decisionmaker--the number of questions to be answered, the time needed for preparing answers, etc. Another source of burden for the decisionmaker is the difficulty of the assessments he or she must make. Questions calling for difficult judgments or which are hard to understand can put considerable stress on a decisionmaker, making the MCDP that much less desirable. Procedures like social judgement theory and the Zionts-Wallenius method make heavy use of holistic assessments--e.g., comparisons or ratings of entire multi-criterion alternatives, rather than only one or two criteria. Holistic assessments are especially difficult and impose a high burden on the decisionmaker.

- Decisionmaker confidence in the procedure: this is a product of a number of factors, some of them subtle. Thus, a decisionmaker's overall confidence in the method may depend on the track record of the procedure (how often it has agreed with the decisionmaker's intuitive feeling about what constitutes the "right" decision), the decisionmaker's understanding and acceptance of the general rationale that underlies the procedure, and perhaps the decisionmaker's general attitude towards analytic vs. "seat-of-the-pants" decisionmaking.
- The meaningfulness of the multi-criterion decision process from the perspective of the decisionmaker. This would include the meaningfulness of the question ,ut to the decisionmaker and the judgments called for, as well as the meaningfulness of the performance criteria themselves. It should be noted that the latter requirement may pose a problem in connection with police use of MCDP's where performance criteria must be formulated so that all are to be maximized or all are to be minimized. Since crime rates will often play a dominant role in assessments of police department performance, one will sometimes be required to formulate all criteria so that, like crime rates, more is worse. But this may mean that some criteria are phrased in very unfamiliar ways--e.g., the percentage of crimes not cleared, the percentage of officers who are dissatisfied with their jobs, etc. This can confuse decisionmakers and make it more difficult for them to supply needed inputs.
- Feasibility and reasonableness of the resources required. In addition to staff and/or decisionmaker time, some MCDP's may require specialized skills not available in the department (making it necessary to hire consultants). MCDP's can also impose other major costs--for computer time, collection of

specialized information, etc. Clearly such costs must be reasonable and must be in balance with the expected benefits of the procedure.

Note that it is not clear whether the intrinsic complexity of a mathematical decision procedure has a major effect on its acceptability to public administrators. Rather complex techniques like the analytic hierarchy process have found widespread acceptance among private sector managers. Furthermore, with the widespread availability of microcomputers and microcomputer packages for some of the procedures described, the computational burden associated with the execution of some of the techniques is becoming less troublesome. On the other hand, if a technique is so complex that the decisionmaker cannot get an intuitive grasp on how--and why--it works, it is unlikely that the decisionmaker will develop the confidence necessary for that multi-criterion decision procedure to become an accepted part of his or her managerial "toolkit."

Promising Multi-Criterion Decision Procedures for Assessing Police Department Performance. The requirements and considerations described above sharply limit the MCDP's potentially acceptable for regular use by police departments in connection with the assessment of overall departmental performance and related questions. Thus, the emphasis on discrete alternatives and the absence of well-established functional relationships between the criteria and the attributes tend to rule out procedures based on (continuous) linear programming formulations. The need to limit the burden on the decisionmaker eliminates approaches such as the Keeney-Raiffa method. Indeed, the dual requirements to limit decisionmaker burden while ensuring that the latter does not feel excluded from the decision process highlights the potential importance of the "mixed" procedures for applications in police and other municipal departments.

When all of the foregoing requirements and considerations are compared to the various multi-criterion decision procedures described previously, four techniques--all of them "mixed" approaches--appear especially promising for regular use by police (and other municipal departments) in connection with performance measurement. These techniques are the analytic hierarchy process, the multi-attribute utility technology (MAUT), compromise programming, and social judgment theory. None of these is a perfect choice; all have advantages and disadvantages. For instance, all of them experience difficulties in handling performance criteria that interact; the problems are especially severe for social judgment theory, since such interactions can substantially weaken and distort the regression results. Nevertheless, all four procedures currently show considerable promise for being practical, useful, and acceptable for regular use in municipal police departments that must come to grips with the need for efficiently and effectively handling multi-criterion performance assessments on a regular basis.

Conclusions and Recommendations

Municipal police departments need help in systematically dealing with the multi-criterion decision and assessment problems they are increasingly--and regularly--facing. Perhaps the most important problem of this type is the assessment of overall police department performance--is the department doing any better now than last month? Last year?

Decision problems of the latter type impose a number of requirements, both mathematical and practical. Consideration of these requirements has reduced the large number of available mathematical multi-criterion decision procedures to four that seem especially promising for the police management context: the analytic hierarchy process, the multi-attribute utility technology, compromise programming, and social judgment theory. While all have both advantages and disadvantages, it is recommended that each approach be adapted to the police context and given careful pilot testing to assess and compare their relative effectiveness, their usefulness, and their acceptance by police managers. While all four techniques potentially have a lot to offer police administrators, they have not received the kind of careful, long-term testing and assessment in the police environment that is needed to judge their actual value and feasibility. An effort to begin such testing would seem to be highly desirable and potentially valuable from the standpoint of improving the ability of police managers to make effective use of the growing policy of good information on departmental performance.

References

1. See, for example, Paul Epstein, Using Performance Measurement in Local Government (New York: Van Nostrand Reinhold Company, 1984); Harry Hatry, Louis Blair, Donald Fisk, John Greiner, John Hall, and Philip Schaenman, How Effective Are Your Community Services? Procedures for Monitoring the Effectiveness of Municipal Services (Washington, D.C.: The Urban Institute and the International City Management Association, 1977); and Gordon P. Whitaker, Stephen Mastrofski, Elinor Ostrom, Roger B. Parks, and Stephen L. Percy, Basic Issues in Police Performance (Washington, D.C.: Department of Justice - National Institute of Justice, July 1982).

2. National Center for Productivity and Quality of Working Life, Total Performance Measurement: Some Pointers for Action (Washington, D.C.: U.S. Government Printing Office, 1978), p. 2.

3. Andrew Van de Ven and Diane Ferry, Measuring and Assessing Organizations (New York: John Wiley and Sons, 1980), pp. 9-19.

4. See, for instance, Whitaker et al., Basic Issues in Police Performance; Hatry et al., How Effective Are Your Community Services, Chapter 6; Harry P. Hatry, "Wrestling with Police Crime Control Productivity Measurement," in Readings on Productivity in Policing, ed. Joan L. Wolfle and John F. Heaphy (Washington, D.C.: The Police Foundation, 1975); Gary T. Marx, "Alternative Measures of Police Performance," Working Paper WP-12-74 (Cambridge, Massachusetts: Massachusetts Institute of Technology, Operations Research Center, October 1974); Michael W. O'Neill, Jerome A. Needle, and Raymond Galvin, "Appraising the Performance of Police Agencies: The PPPM System," Journal of Police Science and Administration, Vol. 8, No. 3 (September 1980), pp. 253-264; and "Modeling Police Agency Performance," ed. Gordon P. Whitaker (Chapel Hill, North Carolina: University of North Carolina, August 1982).

5. James L. Mercer and Susan F. Woolston, Setting Priorities: Three Techniques for Better Decision Making, Management Information Service Report, Vol. 12, No. 9 (Washington, D.C.: International City Management Association, September 1980), and Andre L. Delbecq, Andrew H. Van de Ven, and David H. Gustafson, Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes (Glenview, Illinois: Scott, Foresman and Company, 1975).

6. This description of the nominal group process is adapted from Mercer and Woolston, Setting Priorities, pp. 3-5.

7. See Mercer and Woolston, Setting Priorities, pp. 5-6, and Harold A. Linstone and Murray Turoff, eds., The Delphi Method: Techniques and Applications (Reading, Massachusetts: Addison-Wesley Publishing Company, 1975).

8. This illustration is taken from Mercer and Woolston, Setting Priorities, pp. 5-6, who summarized an example provided in E.S. Quade and W. J. Boucher, eds., Systems Analysis and Policy Planning: Applications in Defense (New York: American Elsevier Publishing Company, 1968), pp. 334-335.

9. There are a great many articles and texts providing reviews and extensive bibliographies of formal MCDP's. See, for example, Erik Johnsen, Studies in Multi-

objective Decision Models (Lund, Sweden: Studentlitteratur, 1968); Milan Zeleny, Multiple Criteria Decision Making (New York: McGraw Hill, 1982); Vira Chankong and Yacov Y. Haimes, Multiobjective Decision Making: Theory and Methodology (New York: Elsevier Science Publishing Company, 1983); Peter C. Fishburn, Decision and Value Theory (New York: John Wiley and Sons, 1964); J. L. Cochrane and M. Zeleny, eds., Multiple Criteria Decision Making (Columbia, S.C.: University of South Carolina Press, 1973)--especially K. R. MacCrimmon, "An Overview of Multiple Objective Decision Making" (pp. 18-44); and Martin K. Starr and Milan Zeleny, eds., Multiple Criteria Decision Making, Studies in the Management Sciences, Vol. 6 (Amsterdam: North Holland Publishing Company, 1977)--especially Martin K. Starr and Milan Zeleny, "MCDM - State and Future of the Arts" (pp. 5-29) and Peter H. Farquhar, "A Survey of Multiattribute Utility Theory and Application" (pp. 59-89). For two very comprehensive bibliographies, see Milan Zeleny, "A Selected Bibliography of Works Related to Multiple Criteria Decision Making," in Cochrane and Zeleny, Multiple Criteria Decision Making, pp. 779-796; and Milan Zeleny, "MCDM Bibliography - 1975," in Milan Zeleny, ed., Multiple Criteria Decision Making: Kyoto, 1975 (New York: Springer-Verlag, 1976), pp. 291-321.

10. See, for instance, H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," in J. Neyman, ed., Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (Berkeley, California: University of California Press, 1951), pp. 481-492; Arthur M. Geoffrion, "Proper Efficiency and the Theory of Vector Maximization," Journal of Mathematical Analysis and Applications, Vol. 22, No. 3 (June 1968), pp. 618-630; and M. Peschal and C. Reidel, "Use of Vector Optimization in Multiobjective Decisionmaking," in Conflicting Objectives in Decisions, ed. David E. Bell, Ralph L. Keeney, and Howard Raiffa (New York: John Wiley and Sons, 1977), pp. 97-122.

11. An alternative A is non-dominated if there exists no other alternative B whose performance is at least as good as that of A for each criterion and whose performance is better than that of A for one or more criteria. Non-dominated points are of interest because the decisionmaker's optimal (most preferred) alternative will be non-dominated. Hence, the set of all non-dominated points will include the decisionmaker's most preferred alternative.

12. See J. P. Evans and R. E. Steuer, "A Revised Simplex Method for Linear Multiple Objective Programs," Mathematical Programming, Vol. 5, No. 1 (1973), pp. 54-72; and Milan Zeleny, Linear Multiobjective Programming (New York: Springer-Verlag, 1974). Alternatives are "continuous" if one or more of the performance criteria can take on any value over a continuous range, with each value (or combination of values) corresponding to a different alternative. Thus, if one possible alternative is characterized by performance criteria $(x_1, x_2, . . . , x_n)$, then $(x_1 + d_1, x_2 + d_2, . . . x_n + d_n)$ is also a possible alternative, where $d_1, . . . , d_n$ can take on any value within their respective ranges (perhaps subject to certain other constraints).

13. Zeleny, Multiple Criteria Decision Making, pp. 248-262; Starr and Zeleny, "MCDM - State and Future of the Arts," p. 23.

14. Arlie L. Bowling and Joseph F. Hair, Jr., "Optimal Decisions on Multiple Objectives Through Canonical Analysis," in James L. Cochrane and Milan Zeleny, eds., Multiple Criteria Decision Making (Columbia, S.C.: University of South Carolina Press, 1973), pp. 729-731; and Fred L. Bookstein, "Statistical Tools for Research into the

U.S. Employment Service," Department of Statistics, University of Michigan (Ann Arbor, September 1978), pp. 34-36.

15. H. J. Einhorn and R. M. Hogarth, "Unit Weighting Schemes for Decisionmaking," Organizational Behavior and Human Performance, Vol. 13 (1975), pp. 171-192; F. L. Schmidt, "The Relative Efficiency of Regression and Simple Unit Predictor Weights in Applied Differential Psychology," Educational and Psychological Measurement, Vol. 31 (1971), pp. 699-714; and Paul J. H. Schoemaker and C. Carter Waid, "An Experimental Comparison of Different Approaches to Determining Weights in Additive Utility Models," Management Science, Vol. 28, No. 2 (February 1982), pp. 182-196.

16. Zeleny, Multiple Criteria Decision Making, pp. 189-197.

17. T. L. Saaty, L. G. Vargas, and R. E. Wendell, "Assessing Attribute Weights by Ratios," Omega, Vol. 11, No. 1 (1983), pp. 9-13.

18. We have, however, eschewed iterative procedures involving lotteries and other probabilistic approaches to the multiple criterion decision problem (although the primary mixed approach of that type - the Keeney-Raiffa procedure - is discussed later). For an example of an iterative interactive version of the Keeney-Raiffa technique, see Rakesh K. Sarin, "Interactive Evaluation and Bound Procedure for Selecting Multi-Attributed Alternatives," in Multiple Criteria Decision Making, ed. M. K. Starr and M. Zeleny, Studies in the Management Sciences, Vol. 6 (Amsterdam: North-Holland Publishing Company, 1977), pp. 211-224.

19. A. M. Geoffrion, J. S. Dyer, and A. Feinberg, "An Interactive Approach for Multi-Criterion Optimization, with an Application to the Operation of an Academic Department," Working Paper No. 176 (Los Angeles, California: Western Management Science Institute, University of California - Los Angeles, July 1971), and Zeleny, Multiple Criteria Decision Making, pp. 361-362.

20. James S. Dyer, "Interactive Goal Programming," Management Science, Vol. 19, No. 1 (September 1972), pp. 62-70.

21. J. Wallenius and S. Zionts, "A Research Project on Multicriterion Decisionmaking," in Conflicting Objectives in Decisions, ed. David E. Bell, Ralph L. Keeney, and Howard Raiffa (New York: John Wiley and Sons, 1977), pp. 76-96; and Zeleny, Multiple Criterion Decision Making, p. 363.

22. Stanley Zionts and Jyrki Wallenius, "An Interactive Multiple Objective Linear Programming Method for a Class of Underlying Non-Linear Utility Functions," Management Science, Vol. 29, No. 5 (May 1983), pp. 519-529.

23. Martin F. Stankard, Jr., Christoph Maier-Rothe, and Shiv K. Gupta, "Choosing Between Multiple Objective Alternatives: A Linear Programming Approach," Project Working Paper (Philadelphia: Management Science Center, University of Pennsylvania, December 1968); and Christoph Maier-Rothe and Martin F. Stankard, Jr., "A Linear Programming Approach to Choosing Between Multi-Objective Alternatives" (Cambridge, Massachusetts: Arthur D. Little, Inc., n.d.).

24. Ralph E. Steuer, "An Interactive Multiple Objective Linear Programming Procedure," in Multiple Criteria Decision Making, ed. Martin K. Starr and Milan Zeleny,

Studies in the Management Sciences, Vol. 6 (Amsterdam: North-Holland Publishing Company, 1977), pp. 225-239; and Zeleny, *Multiple Criteria Decision Making*, p. 364.

25. Steuer, "An Interactive Multiple Objective Linear Programming Procedure," p. 232.

26. See Zeleny, *Multiple Criteria Decision Making*, Chapters 5, 6, and 10.

27. R. Benayoun, J. Tergny, and D. Keuneman, "Mathematical Programming with Multi-Objective Functions: A Solution by P.O.P. (Progressive Orientation Procedure)," *Metra*, Vol. 9, No. 2 (1970), pp. 279-299; R. Benayoun, J. De Montgolfier, J. Tergny, and O. Larichev, "Linear Programming with Multiple Objective Functions: Step Method (STEM)," *Mathematical Programming*, Vol. 1, No. 3 (1971), pp. 366-375; and Bernard Roy, "Problems and Methods with Multiple Objective Functions," Seventh Mathematical Programming Symposium (The Hague, September 14-18, 1970).

28. Among the many good references on goal programming, see Yuji Ijiri, *Management Goals and Accounting for Control* (Chicago, Illinois: Rand-McNally, 1965); Sang M. Lee, "Goal Programming for Decision Analysis of Multiple Objectives," *Sloan Management Review*, Vol. 14 (1973), pp. 11-24; Sang M. Lee and Richard L. Morris, "Integer Goal Programming Methods," in Martin K. Starr and Milan Zeleny, eds., *Multiple Criteria Decision Making* (Amsterdam: North-Holland Publishing Company, 1977), pp. 273-289; and James S. Dyer, "Interactive Goal Programming."

29. Ralph E. Steuer, "Multiple Objective Linear Programming with Interval Criterion Weights," *Management Science*, Vol. 23, No. 3 (November, 1976), pp. 305-316; and Zeleny, *Multiple Criteria Decision Making*, p. 364.

30. See the discussions and references given in Schoemaker and Waid, "An Experimental Comparison," p. 184; Paul J. H. Schoemaker, "Behavioral Issues in Multiattribute Utility Modeling and Decision Analysis," in *Organizations: Multiple Agents with Multiple Criteria*, ed. J. Morse (New York: Springer-Verlag, 1981), pp. 447-464 (especially p. 447); and Benjamin F. Hobbs, "A Comparison of Weighting Methods in Power Plant Siting," *Decision Sciences*, Vol. 11, No. 4 (October 1980), pp. 725-737 (especially p. 728).

31. Ralph L. Keeney and Howard Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* (New York: John Wiley and Sons, 1976).

32. These versions of the preference independence assumptions are based on the discussion in R. K. Sarin and R. L. Winkler, "Performance-Based Incentive Plans," *Management Science*, Vol. 26, No. 11 (November 1980), pp. 1131-1144.

33. See Keeney and Raiffa, *Decisions with Multiple Objectives*, pp. 299-300, and James S. Dyer and Rakesh K. Sarin, "Measurable Multiattribute Value Functions," *Operations Research*, Vol. 27, No. 4 (July-August 1979), pp. 810-822.

34. Starr and Zeleny, "MCDM - State and Future of the Arts," pp. 20-21.

35. Among the numerous publications on MAUT, see Ward Edwards and J. Robert Newman, *Multiattribute Evaluation*, Sage Quantitative Applications in the Social Sciences, No. 26 (Beverly Hills, California: Sage Publications, 1982); Ward Edwards, "How to Use Multiattribute Utility Measurement for Social Decisionmaking," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-7, No. 5 (May 1977), pp. 326-340; Ward Edwards,

Marcia Guttentag, and Kurt Snapper, "A Decision-Theoretic Approach to Evaluation Research," in Handbook of Evaluation Research, Vol. 1, ed. Elmer L. Struening and Marcia Guttentag (Beverly Hills, California: Sage Publications, 1975), pp. 139-181; and Ward Edwards, "Multiattribute Utility for Evaluation: Structures, Uses, and Problems," in Handbook of Criminal Justice Evaluation, ed. Malcolm W. Klein and Katherine S. Teilman (Beverly Hills, California: Sage Publications, 1980), pp. 177-215.

36. The basic theory and numerous applications can be found in Thomas L. Saaty, "A Scaling Method for Priorities in Hierarchical Structures," Journal of Mathematical Psychology, Vol. 15 (1977), pp. 234-281; Thomas L. Saaty, The Analytic Hierarchy Process (New York: McGraw Hill, 1980); Thomas L. Saaty and Luis G. Vargas, The Logic of Priorities (Boston: Kluwer-Nijhoff Publishing, 1982); and Thomas L. Saaty, Decision Making for Leaders (Belmont, California: Lifetime Learning Publications, 1982). For some critiques of the process, see S.R. Watson and A.N.S. Freeling, "Assessing Attribute Weights," Omega, Vol. 10, No. 6 (1982), pp. 582-583; Valerie Belton and Tony Gear, "On a Short-Coming of Saaty's Method of Analytic Hierarchies," Omega, Vol. 11, No. 3 (1983), pp. 228-230; and S.R. Watson and A.N.S. Freeling, "Comment on: Assessing Attribute Weights by Ratios," Omega, Vol. 11, No. 1 (1983), p. 13.

37. See, for instance, Kenneth R. Hammond and Thomas R. Steward, "Social Judgment Theory," in Human Judgment and Decision Processes, ed. Martin F. Kaplin (New York: Academic Press, 1975), pp. 271-312; Kenneth R. Hammond, "Externalizing the Parameters of Quasi-rational Thought," in Multiple Criteria Decision Making: Kyoto 1975, ed. M. Zeleny, Lecture Notes in Economics and Mathematical Systems (New York: Springer-Verlag, 1976), pp. 75-96; Kenneth R. Hammond, John Rohrbaugh, Jeryl Mumpower, and Leonard Adelman, "Social Judgment Theory: Applications in Policy Formation," in Human Judgment and Decision Processes in Applied Settings, ed. Martin F. Kaplan and Steven Schwartz (New York: Academic Press, 1977), pp. 1-29; Kenneth R. Hammond, Richard L. Cook, and Leonard Adelman, "POLICY: An Aid for Decision Making and International Communication," Columbia Journal of Work Business (Fall 1977), pp. 79-93; Thomas R. Steward and Linda Gelberd, "Analysis of Judgment Policy: A New Approach for Citizen Participation," American Institute of Planners Journal, Vol. 42 (1976), pp. 33-41; and John Rohrbaugh and Paul Wehr, "Judgment Analysis in Policy Formation: A New Method for Improving Public Participation," Public Opinion Quarterly, Vol. 42 (Winter 1978), pp. 521-532.

38. Many of these applications have been drawn from a recent report by Gloria A. Grizzle, "Relative Importance of Different Performance Dimensions," Working Paper 82-8 (Tallahassee, Florida: The Osprey Company, November 1982).

39. Samuel E. Bodily, "Merging the Preferences of Interest Groups of Efficiency and Equality of Service in the Design of Police Sectors," in Police Deployment, ed. Richard Larson (New York: Lexington Books, 1978), pp. 103-122.

40. Kurt J. Snapper and Donald A. Seaver, "Application of Decision Analysis to Program Planning and Evaluation," Technical Report (Falls Church, Virginia: Decision Science Consortium, n.d.).

41. Edwards, "Multiattribute Utility for Evaluation," pp. 208-209.

42. Hammond, "Externalizing the Parameters of Quasi-rational Thought," pp. 89-92.

43. Grizzle, "Relative Importance of Different Performance Dimensions."

44. Ralph L. Keeney and Keshavan Nair, "Selecting Nuclear Power Plant Sites in the Pacific Northwest Using Decision Analysis," in Conflicting Objectives in Decisions, ed. David E. Bell, Ralph L. Keeney, and Howard Raiffa (New York: John Wiley and Sons, 1977), pp. 298-322.

45. Ralph L. Keeney, "A Decision Analysis with Multiple Objectives: the Mexico City Airport," Bell Journal of Economics and Management Science, Vol. 4, No. 1 (1973), pp. 101-117.

46. Howard M. Ellis and Ralph L. Keeney, "A Rational Approach for Government Decisions Concerning Air Pollution," in Analysis of Public Systems, ed. A. W. Drake, R. L. Keeney, and P. M. Morse (Cambridge, Massachusetts: MIT Press, 1972), pp. 376-400.

47. William T. Stewart, "A 'Yardstick' for Measuring Productivity," Industrial Engineering (February 1978), pp. 34-37.

48. Jeff Karako and Harvey Wolf, "Evaluating a Street Department's Performance," Public Productivity Review, Vol. 6, No. 1-2 (1982), pp. 122-127.

49. Edwards, "Multiattribute Utility for Evaluation," pp. 207-208.

50. Saaty, The Analytic Hierarchy Process.

51. John Rohrbaugh and Robert Quinn, "Evaluating the Performance of Public Organizations: A Method for Developing a Single Index," Journal of Health and Human Resource Administration, Vol. 2 (1980), pp. 343-354.

52. Hammond et al., "Social Judgment Theory: Applications in Policy Formation."

53. Zeleny, Multiple Criteria Decision Making, pp. 336-338.