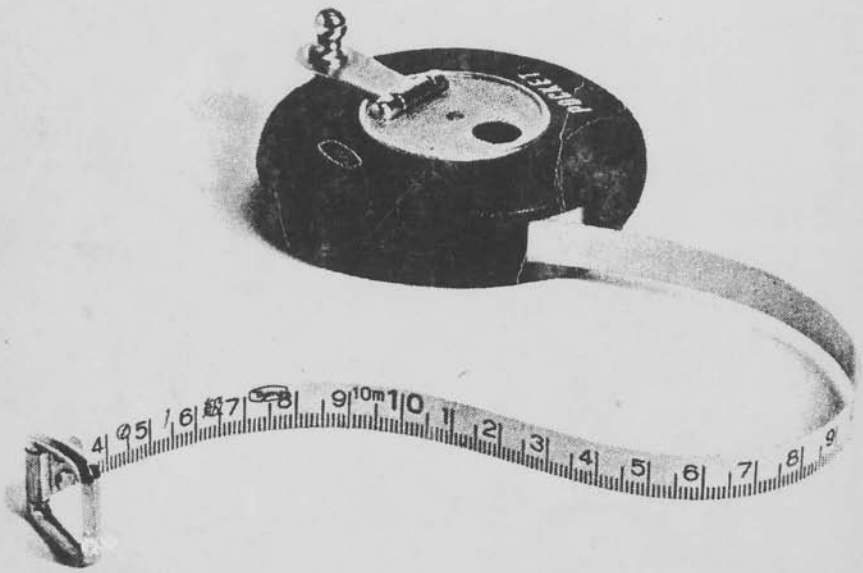
EVALUATION
OF COMPLEX
POLICY PROBLEMS



edited by ANDREAS FALUDI
HENK VOOGD

DELFTSCHE UITGEVERS
MAATSCHAPPIJ B.V.

EVALUATION OF COMPLEX POLICY PROBLEMS

EVALUATION OF COMPLEX POLICY PROBLEMS

edited by

ANDREAS FALUDI
University of Amsterdam

and

HENK VOOGD
Delft University of Technology

Delftsche Uitgevers Maatschappij B.V. - 1985

Delftsche Uitgevers Maatschappij
Koornmarkt 70
2611 EJ Delft
THE NETHERLANDS
phone: (015) 123725

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Evaluation

Evaluation of complex policy problems / ed. by A. Faludi and Henk Voogd. - Delft : Delftsche U.M. - Ill.
ISBN 90-6562-071-0
SISO 719.12 UDC 711(492)
Trefw.: planologie.

CONTENT

PREFACE

Evaluation is a major aspect of contemporary public planning and policy-making. The range of methods available nowadays is vast and reaches from the rough-and-ready measurement of costs and benefits and the use of checklists to complex multi-criteria analyses and optimization techniques. But so far insufficient attention has been paid to the applicability of these methods in practice, and this despite the fact practical applications should hold many lessons for those concerned with the development of planning methodology in general and the methodology of evaluation in particular. This methodology of evaluation - broadly conceived - is what this book is about.

The papers in this volume all focus on (**ex-ante**) evaluation of plans or projects. Plan or project evaluation forms part of a broader approach to rendering complex public choices more justifiable. Usually, though, the evaluation methods which are available are insufficient for solving all the many practical problems encountered in practice. The authors of the papers included in this volume have attempted to bridge the theory-practice gap. Their papers cover a range of aspects and reflect the state of the art in evaluation. Care has been taken to cover practical applications alongside with conceptual advances. The papers allow several critical conclusions to be drawn and stimulate new ideas which should assist with the further development of evaluation in public planning and policy-making.

This book originates from an international workshop held in 1984 at Delft in The Netherlands. The participants were practitioners working for government departments as well as academic researchers and consultants. Being experts in the field, they come from various disciplinary backgrounds and represent a number of nationalities. They engaged in fruitful and stimulating discussions. That this led to a real cross-fertilization of ideas became evident when many authors felt it apposite to substantially modify their papers for inclusion in this book.

Several people gave expert assistance with the preparation of this volume. Our special thanks go to Ir. Mark Maimone. He not only made notes during the discussions at the workshop which proved invaluable, he also delicately performed the thankless task of editing papers written by authors whose misfortune it is not to have English as their mother tongue. The task of subsequently word-processing the majority of the papers in this volume has been in the care of Mrs. Lidy Verbeek. She has discharged her responsibilities in the meticulous way which we have come to expect of her.

Delft, July 1985

Andreas Faludi
Henk Voogd

Acknowledgement

The workshop and this book have been made possible by financial support of the Dutch National Science Foundation (ZWO), the Ministry of Finance and the Vakgroep Civiele Planologie of the Delft University of Technology.

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EVALUATION OF COMPLEX POLICY PROBLEMS
SOME INTRODUCTORY REMARKS

A. Faludi
H. Voogd

INTRODUCTION

In public policy making, more and more attention goes to the explicit analysis of decision problems. Undoubtedly, the present economic crisis forms an impetus for this. It makes the importance of rigorous analyses of plans and policies much more plausible.

One of the most profound challenges which planners face nowadays is a growing sense of uncertainty: about society, the economy, resources, the environment, and so forth. Significant demographic and technological changes occur alongside with changes in the economy and in political style and preferences. Will the gloomy period of economic recession, or stagnation, be followed by recovery? Will the current shift towards more right-wing politics be followed by yet another swing of the pendulum to left? Will the microchip lead to changes in public planning and policy making?

Uncertainties notwithstanding, policies must be developed and decisions made. Many methods and techniques are available to assist the decison taker. An important cluster of such methods relates to **evaluation.** These are methods which enable planners, or policy analysts, to structure information concerning possible alternative choices. There seems to be no general consensus, though, about the most appropriate method of evaluating alternatives. Worse still, aggreement is lacking even about preliminaries, for instance whether explicit and systematic evaluation methods make valid contributions to policy choices.

The prime purpose of this book "Evaluation of Complex Policy Problems" is to discuss the major constraints on evaluation and to give insight into new directions in policy evaluation.

THE SCOPE OF THE BOOK

The purpose of evaluation is to give as objective a description of the situation in question as is humanly possible, and to appraise it prior to making a considered choice. As this term is used here, "situation" includes alternative options available to the decision taker. Options can refer to elaborately developed alternative plans or projects, but also to preliminary ideas between which a choice must be made at an early stage of the preparation of a policy. Since options always belong to a decision taker, it follows that one decision taker's situation will be different from that of another decision taker. So, where several decision takers operate in one and the same environment, we can distinguish several "situations" - in fact one per decision taker. That these situations are interdependent in the sense of

one decision taker's choice influencing those of others only adds to the uncertainty under which choices are being made.

Various kinds of evaluation exist. This volume focuses primarily on so-called **ex-ante** evaluation, viz. structuring information concerning possible alternatives prior to making a choice between them. Presently, a broad spectrum of **ex-ante** evaluation methods is available. They may be classified roughly into three broad categories:

*** Monetary Evaluation Methods**
These methods are designed to provide a framework for a monetary assessment of alternative plans or projects. Examples are cost-benefit analysis (see Dasgupta and Pearce, 1972; Mishan, 1973), cost-effectiveness analysis (see English, 1968; Meike, 1973) and threshold analysis (see Kozlowski, 1968; Malisz, 1970).

*** Overview Methods**
These methods present not only monetary but also non-monetary information concerning alternatives and their potential effects. These are expressed in the form of schemes or tables. Well-known examples are the planning balance sheet method (Lichfield, 1968; Lichfield c.s., 1975) and the score card method (Rand Corporation, 1977).

*** Multi-criteria Methods**
These methods add an extra dimension. They enable the analysis and presentation of available information concerning alternatives in the light of multiple and heterogeneous criteria. Depending on the type of information used, a further distinction can be drawn within this category between so-called "quantitative" methods, like goals-achievement analysis (Hill, 1973) and concordance analysis (Van Delft and Nijkamp, 1977), "qualitative" methods, like permutation analysis (Paelinck, 1978) and regime analysis (Hinloopen c.s., 1983) and "mixed data" methods (Voogd, 1983).

Almost all research efforts in the fifties were devoted to the development of monetary (**ex ante**) evaluation. The fact that there is usually more than one criterion of choice was conveniently forgotten. This changed in the late sixties, due to substantial revisions that took place in the field of planning and policy-making (see for instance: Faludi, 1984, first published 1973). Planners and politicians became increasingly aware of the need for a more integrated approach to economic, environmental, social and physical developments. In its wake, planners became increasingly aware of the fact that the objectives pertaining to these various fields did not lend themselves to being expressed in one or two performance criteria. Rather, objectives differ, frequently to the point where there is outright conflict between them. This is particularly true where plans or policies are designed to increase allocative efficiency and social equity and deal with environmental impact at the same time.

These changes in policy objectives apart, the involvement of interest groups has also changed. As a consequence, the focus of planning and policy-making has shifted towards greater concern with the decision-making process and the equity of decisions. So there is a greater involvement now of the public at large in decisions which are likely to have a direct or indirect effect on them. Principally, this occurs through consultation and participation.

This development has had a definite impact on research into evaluation and the way evaluation methods - such as those mentioned above - are used. Instead of searching for "optimal" solutions in terms of monetary costs and benefits, at

present attention is directed more and more towards the systematic analysis of a variety of options in the light of - sometimes conflicting - objectives, the aim being to provide a broad basis for justifying decisions. In so doing, **accountability** to a range of various interests becomes a dominant concern in contemporary **ex-ante** evaluation.

A wide variety of publications notwithstanding, most attention in the literature still goes to the more "technical" aspects of evaluation. So academic discourse concerning **ex-ante** evaluation tends to concern the methods and techniques as such, in most instances leaving the practical problems encountered in evaluation and their consequences out of consideration. The workshop on which this volume draws has given special emphasis, therefore, to issues and practical constraints encountered in **ex-ante** evaluation in practice. In particular, a link has been forged with the well-known "strategic choice approach" to planning (see Friend and Jessop, 1977, 1st edition 1969; Hickling, 1974; Sutton c.s., 1977; Faludi and Mastop, 1982). This approach provides a general framework for public planning, and also includes several techniques adapted to that framework.

THE STRUCTURE OF THE BOOK

The structure of this book is simple. More conceptual chapters in Part I are followed by Dutch case studies of the application of evaluation methods in Part II. But, of course, the distinction is far from rigid. The authors of the conceptual chapters base themselves on a wide range of practical experiences, and many comments in the case studies reach beyond the confines of the immediate problem at hand.

That all case studies are Dutch is no accident but reflects an aim of the workshop which was to confront the state-of-the-art in **ex-ante** evaluation in The Netherlands with a wide range of international experiences. That is also the reason why the papers in Part I on **Evolving Approaches to Evaluation** are predominantly written by authors coming from abroad. Thus, M. Hill explores a theme which has been discussed in the planning literature for some considerable time now: the relationship between planning and its context. He offers a taxonomy of decision-making situations and relates it to evaluation methods. In focusing on distributional effects, D. H. Miller adds an important dimension to the discussion of these methods. Usually, evaluation is most concerned with the effectiveness and efficiency of proposed courses of action. N. Lichfield broadens the scope of planning balance sheet analysis which he has turned into such a fine art into impact evaluation, like Hill relating evaluation to its context at the same time.

The exceptions to the rule of conceptual papers in Part I having been written by authors from outside The Netherlands are the ones by J. Buit, P. Nijkamp and P. Daru. In his paper, Buit reflects upon evaluation in a situation which, although making itself felt worldwide, is particularly evident in The Netherlands: the dramatic change from the expectation of growth and abundance to a perspective reflecting the experience of decline and scarcity. In his discussion of a more technical issue: the relationship of evaluation and information - in particular of a spatial kind - Nijkamp draws on international surveys. The paper by Daru similarly concentrates on one specific aspect of evaluation: the communication of its results by graphic means of presentation.

This is an important issue in overcoming a problem which the workshop returned to frequently in its deliberations: closing the gap between technical analyses on the one hand and "day-to-day" decision-making on the other. This has been a long-

standing concern also of all those concerned with he development of the so-called strategic choice approach mentioned above. The next two papers relate to this approach. P. Dello gives an exposition of what it involves and pays specific attention to evaluation as an all-pervasive aspect of working with strategic choice. A. Hickling, who has considerably refined strategic choice from its early beginnings in Friend and Jessop (1977, 1st edition 1969), readily agrees that "it is impossible to make progress without making judgements all the time". He distinguishes various modes of work within strategic choice and, with his usual skill in giving diagrammatic expression to his idea, proposes a "five-finger model" of evaluation.

Part II with the **Dutch applications** opens with one that relates to the application of strategic choice to identifying suitable sites for an LPG-terminal in The Netherlands by R. van de Graaf. A. Sorber reports from his central vantage point of the Dutch Ministry of Finance on the practice of project appraisal in Dutch central government. The paper by J. van Staalduine concerns the same level of government, but relates to physical planning, a field in which the Dutch have received some considerable international acclaim for their efforts in the past. J. W. de Vos gives a detailed account of the making of a regional plan in one of the Dutch provinces, whereas the remaining two papers by S. J. Bennema **et al.** respectively M. Maimone concern individual infrastructure projects which, in a densily populated country like The Netherlands, are of considerable complexity. Affecting, as they do, many groups with conflicting interests, they are ideally suited for demonstrating the potential of multi-criteria evaluation.

The volume concludes with comments by the editors of this volume who were at the same time responsible for convening what turned out to be a very congenial international group of experts in the field of evaluation. It is hoped that some of the inspiration which they received during the discussions comes through in those last pages.

REFERENCES

Dasgupta, A.K. and D.W. Pearce, **Cost-Benefit Analysis: Theory and Practice**, MacMillan Press, Londen (1972)

English, M.J. (ed.), **Cost Effectiveness**, Wiley, New York (1968)

Faludi, A. **Planning Theory**, Pergamon, Oxford (1984) (first published 1973)

Faludi, A. and J.H. Mastop, The 'I.O.R.-School' - The development of a planning methodology, **Environment and Planning B** (1982), 9, 241-256

Friend, J.K. and W.N. Jessop, **Local Government and Strategic Choice**, Pergamon, Oxford (1977) (1st edition: Tavistock, London, 1969)

Hickling, A., **Managing Decisions**, Mantec Publications, Rugby (1974)

Hill, M., **Planning for Multiple Objectives**, Regional Science Research Institute, Monograph Series 5, Philadelphia (1973)

Hinloopen, E., P. Nijkamp and P. Rietveld, Qualitative Discrete Multiple Criteria

Choice Models in Regional Planning, **Regional Science and Urban Economics** (1983) 77-102

Keeney, R.L. and H. Raiffa, **Decisions with Multiple Objectives; Preferences and Value Tradeoffs**, Wiley, New York (1976)

Kozlowski, J., Threshold Theory and the Sub-regional Plan, **Town Planning Review**, July (1968) 15-34

Lichfield, N., Economics in Town Planning, **Town Planning Review**, vol. 39 (1968) 79-96

Lichfield, N., P. Kettle and M. Whitbread, **Evaluation in the Planning Process**, Pergamon Press, Oxford (1975)

Malisz, B., Direct Application of Threshold Analysis to Regional Planning, **Papers from the Seminar on Threshold Analysis**, C.E.S. London (1970)

Meike, U., **Cost effectiveness analysis als Planungsinstrument**, Vandenhoeck & Ruprecht, Göttingen (1973)

Mishan, E.J., **Cost-Benefit Analysis, an informal introduction**, Allen & Unwin, London (1974)

Nijkamp, P., **Environmental Policy Analysis**, Wiley, London (1980)

Paelinck, J.H.P., Qualiflex, A Flexible Multiple Criteria Method, **Economic Letters**, vol. 3 (1978) 193-197

Rand Corporation, Protecting an Estuary from Floods - A Policy Analysis of the Oosterschelde (prepared for the Netherlands Rijkswaterstaat)/B.F. Goeller c.s., Vol. I-VI, Rand Corporation, Santa Monica (1977)

Rietveld, P., **Multiple Objective Decision Methods and Regional Planning**, North Holland Publ., Amsterdam (1980)

Sutton, A., A. Hickling, and J.K. Friend, **The analysis of policy options in structure plan preparation: the strategic choice approach.** Institute for Operational Research, Tavistock, London (1977)

Van Delft, A. and P. Nijkamp, **Multicriteria Analysis and Regional Decision-Making**, Martinus Nijhoff, The Hague (1976)

Voogd, H., **Multicriteria Evaluation for Urban and Regional Planning**, Pion, London (1983)

PART ONE: EVOLVING APPROACHES IN EVALUATION

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

DECISION-MAKING CONTEXTS AND STRATEGIES FOR EVALUATION

M. Hill
Technion-Israel Institute of Technology
Dept. of Urban Planning

INTRODUCTION

Methods for the evaluation of public policy alternatives generally assume a standard context, as if the institutions that are charged with the responsibility for public policy-making are always cast in the same mould. The usual simplifying assumption is that the evaluation is intended to serve a central decision-making body acting according to a unitary perception of the public interest. This decision-making body is assumed to be empowered to decide on that course of action which, in its judgement, best serves the public interest. Further, this wise and rational body is able to follow through by implementing the policies that been adopted. Thus, the costs and the benefits 'to whomsoever they accrue' of the policy alternatives are indentified, measured, aggregated and compared, thereby indicating the preferred course of action and the 'devil take the hindmost'. In a world in which all outcomes can be predicted with certainty or in which the extent of uncertainty (and hence the risk of a wrong decision) is known, the rational decision is therefore obvious.

However, we have come to learn that the real world of planning and policy-making is a far cry fom these simplistic assumptions. Administrative bureaucracies and political decicion-making bodies can be quite varied in the axtent of their power and control, perception of their responsibility, extent to which they are able to act in accordance with the public consensus, extent of their accountability, and so forth. It therefore appears rash to assume away all this complexity in order to adopt a pure formula for arriving at an optimal decision in the public interest. Neither, for that matter, does it seem valid to succumb to the complexity of the varied decision-making contexts and plead the impossibility of pursuing a course of action which is rationally arrived at. On the contrary, it seems essential to take note of the complexity and try to take it into account in developing approaches and tools for the evaluation of policy alternatives.

In this paper we set out to demonstrate that the mode of decision-making has important implications for the choice of the evaluation strategy. An evaluation strategy can be analyzed in terms of a set of evaluation variables. These evaluation variables can be pointers to the choice of the evaluation methodology which is appropriate for the particular decision-making situation.

In the first section of the paper the evaluation variables are reviewed and the relationship of thes variables to various planning methodologies is indicated. In the following section several decision-making modes are identified and analyzed in terms of a set of decision varables. In the final section of the paper the implications of the decision variables for the evaluation variables, and hence the evaluation strategy are postulated.

EVALUATION VARIABLES

Among the evaluation variables which may be affected by the decision-making context are the following:

- The perception of the public interest
- The treatment of uncertainty
- The number of stages in the evaluation process
- The assessment of time preference
- Ex ante, continuous or ex post evaluation
- Distributional equity
- Comprehensibility (transparency) of methodology
- Sophistication of evaluation procedures
- Extent that effects on all interested parties are recorded
- Optimum-seeking or satisficing evaluation method
- Comprehensiveness or disjointedness of evaluation procedure
- Interactive nature of evaluation process.

These evaluation varables vary in accordance with the methodology that is employed as we shall now demonstrate.

The perception of the public interest.
Cost-benefit analysis (Peskin and Seskin, 1975) purports to measure allocative efficiency and implies a unitary public interest, i.e., that these ends pertain equally to all members of society (Meyerson and Banfield, 1955). Even when multiple goals are considered the analyst may take a unitary view of the public interest (Major, 1977). On the other hand, evaluation methods such as the planning balance sheet (Lichfield 1966, 1975) and the goals-achievement matrix (Hill, 1968, 1973) assume that the public interest is a composite of multiple interests. (1)

The treatment of uncertainty.
Cost-benefit analysts have for a long time employed analytical techiques for the analysis of the uncertainty associated with future courses of action (Dasgupta and Pearce, 1972) but the proponents of multiple objective techniques have not always tried to tackle uncertainty along with other complexities. However, Nijkamp and Van Delft (1977) have incorporated a stochastic element into concordance analysis in order to treat uncertainty about the impact of plans and uncertainty of future preferences. Voogd (1980) has developed a stochastic geometric scaling procedure while Stuart (1974) uses sensitivity analysis in order to relate to uncertainty about impacts, costs and preferences.

The number of stages in the evaluation study.
As cost-benefit analysts moved away from a narrow efficiency criterion perspective they broadened the analysis by means of lexicographic ordering (Steiner, 1977). Only projects which have positive net benefits according to cost-benefit analysis are acceptable. Only then are these projects evaluated in terms of other objectives. Recently, there has been work done on multistage decision-making processes in multiple objective evaluation contexts. In the recent volume by Nijkamp and Spronk (1981) there are three papers addressed to this problem.

Treatment to time preference.
The treatment of time preference has been a central feature of cost-benefit analysis (Mishan, 1971). It is less common in multiple objective evaluation methodologies. It is particularly pertinent in the case of capital budgeting and financial planning and Spronk (1981) demonstrates how this may be taken into con-

sideration in a multiple goal evaluation procedure.

Ex-ante, ex-post and continuous evaluation.
Methods of evaluation such as cost-benefit analysis, cost-effectiveness analysis, the planning balance sheet, the goals-achievement matrix and other multiple criteria methods were generally developed for purposes of ex-ante evaluation. Developed primarily by people coming from the fields of economics, operations research and urban planning, these methods of evaluation were primarily conceived for purposes of enhancing decision-making about future courses of action. By contrast, ex-post evaluation for the assessment of the effectiveness of public progams has been largely the province of sociologists and psychologists (Rossi and Freeman, 1982) with a relatively small involvement of economists (Cain and Holloster, 1977). This has recently neem supplemented by the analysis of the implementation of public programs influenced by political scientists (Alterman, 1982, 1983). The present challenge is the development of continuous evaluation tools that can serve ongoing decision-making oriented both to short-term periodic decisions and longer term strategic decisions (Alterman, Carmon and Hill, 1984).

Distributional equity.
Whereas cost-benefit analysis explicitly excludes distributional effects, starting in the late 1960's several economists proposed that it be supplemented by the analysis of the distribution of benefits in order to facilitate decision-making concerning distributional equity (Marglin 1967, Weisbrod, 1977). The planning balance sheet and the goals-achievement matrix array the effect of alternative courses of action on various publics, thus facilitating an analysis of distributional equity. Miller (this volume) and Schermer (1975) employ multiple objective frameworks in order to trace distributional effects.

Comprehensibility (transparency) of the evaluation methodology.
Much evaluation work is beyond the comprehension of lay decision-making bodies. The more complex the methodology used, the more difficulty the non-expert has in comprehending. The more aggregate the measures of the outcomes (as in cost-benefit analysis), the more difficulty interested parties may have in determining how well off they will be as a result of the plan under consideration. The multiple objective, multiple interest, balance sheet methods such as the planning balance sheet, the goals achievement matrix, the net/benefit assessment process developed by the American Institute of Planners for the city of Simi Valley (Schamberg 1977) and Poulton's land use evaluation matrix (1981) provide the greatest transparency, and hence accountability, to the affected parties. It is worthy of note that the U.K. Standing Advisory Committee on Trunk Road Assessment, chaired by Sir George Leitch, recommended a comprehensive framework for appraisal (1979) on the above lines with particular emphasis on its comprehensibility.

Level of sophistication of evaluation procedure.
The level of sophistication in mathematical terms of the evaluation methodology is negatively correlated with the previous variable, the level of comprehensibility. The Leitch committee explicitly rejected multicriteria weighting techniques and concordance analysis since (P.6) the techniques appear to reduce the capacity for judgement (of the decision-making body) and they "tend to be complex and thus inappropriate for day to day use and are unlikely to command public confidence." The highly sophisticated mathematically-based multiple objective techniques developed by those engaged in operations research (Zeleny, 1975 and Cochrane and Zeleny, 1973), are well beyond the comprehension of lay public decision-making bodies.

Effects on all interested parties.
The extent that effects on all interested parties are recorded is not really a separate and independent variable and is subsumed under the transparency variable. We have treated it separately because of the important function that this specific information may fulfil in determining the acceptability and ultimate implementation of the plan. If those in the decision-making body see that they are favorably affected by the plan, the chances of its implementation are enhanced. Whereas, if they see that they are unfavorably affected, the acceptability of the plan is significantly reduced. It is clear that balance sheet and matrix approaches such as those mentioned above help to trace the effects on all interested parties.

Optimumseeking or satisficing techniques.
Cost-benefit analysis which is oriented to optimizing the efficiency criterion is clearly an optimum-seeking technique. Multiple objective mathematical models of the operations research genre are frequently optimum seeking (Sengupta et al., 1973). By contrast, several analysts have recognized that in the face of multiple and frequently conflicting interests, the best strategy might be a satisficing one. This has recently been employed in several methodological innovations by Nijkamp and Vos (1977), by Lomovasky and Hill (1984) and by Werczberger (1983).

Comprehensiveness or disjointedness of evaluation procedure.
Evaluation techniques range from those methods which attempt to trace all the effects of the course of action under consideration, in a comprehensive manner, to those methods which explicitly narrow their range of consideration of effects. Among the more comprehensive approaches are the various balance sheet and matrix approaches mentioned above. Cost-benefit analysis, which tries to trace all the effects of the course of action which can be expressed in quantitative, preferable monetary terms, is a special case among the comprehensive procedures. Among the more narrowly focussed (disjointed) techniques are threshold analysis (Koslowski and Hughes, 1972) and cost-effectiveness analysis (Goldman, 1967 and Levin, 1975).

Interactive nature of the evaluation process.
Most evaluation techniques do not have an interactive component. However, faced with the desire to involve affected and interested citizens in the evaluation process in an interactive manner, there have been some innovative techniques. Among the analysts who have attempted to introduce interactive components in an evaluation methodology are Stuart (1974), Lomovasky and Hill (1984) and Werczberger (1983).

MODES OF DECISION-MAKING

It is obvious that bureaucratic structures and political institutions are not divorced form the social and economic context in which they operate. Decentralized decision-making is a sine qua non of a socio-economic structure in which economic activity is diffuse, based on private enterprise or on the cooperative ownership of each enterprise by its operatives. On the other hand, authoritarian societies and their governments, with centralized economic control, do not go along with a decision-making structure in which power over public policy is dispersed and shared with the citizenry. It is certainly a valuable and worthwhile enterprise to analyze the relationship between the social and economic context and the institutional and bureaucratic structures and consequent decision-making processes. However, this is not our thrust and we shall not pursue this direction any further at this stage. Instead we shall assume alternative modes of decision-making and planning while relating them to varying political-institutional struc-

tures and processes.

As a point of departure we shall adopt **John Friedmann's four types** of planning and decision-making modes or styles (Friedmann, 1973):

(1) Command or centralized planning.
(2) Policies planning.
(3) Corporate planning.
(4) Participatory planning.

There are no doubt other possible modes which can be derived independently. These four modes will, however, serve for presenting an argument for varying the evaluation strategy in accordance with the pattern of decision-making. In our subsequent discussion we shall deal in greater detail with the implications of the planning and decision-making modes for evaluation procedures. At this stage, short definitions of the various styles of planning will suffice.

Command or central planning occurs under conditions when power is highly centralized. This is the planning model which the evaluation literature generally assumes to occur. The predominant means of control is by sanctions in order to meet the required objectives. We shall distinguish between two types of control. In the **command-initiatoty mode,** control is achieved by budget allocations whereas in the **command-regulative mode** the central planning and decision-making bodies have statutory control, as in the case of statutory master-plans which are legally binding.

Policies planning is characteristic of **weakly** centralized governmental systems, as in the case of a hierarchical governmental structure without coercive power. Control is achieved by the central decision-making body through the provision of guidelines and decision-criteria for the subordinate levels but without being able to enforce these. Instead, material incentives by central goverment encourage others to follow the guidelines. Also, information is disseminated to ensure that the various bodies operate from the same information base, thus reducing uncertainty on this account.

Corporate planning occurs when power is **decentralized among a small number of** corporate bodies such as trade unions, industrial and commercial conglomerates, farmers'organizations, the church and universities. Where power is so organized, each major corporate body has an effective veto with respect to any action affecting its own area of influence. No policy can be adopted without the consent of those bodies that will be directly affected. Decisions are arrived at as a result of a negotiating or bargaining process. Control is achieved by means of a normative compliance with agreements reached through bargaining and negotiation.
Participatory planning occurs when power is dispersed among many actors and resides in various types of social organizations such as a neighborhood association, the town meeting and workers institutions such as cooperatives or agricultural settlements. The predominant method of control is by voluntary compliance of participants as a result of group deliberation. Corporate bodies are usually controlled from the top, whereas community groups usually share in all important decisions concerning use of their resources.

Although in any given situation a particular decision-making style or mode may be dominant, it is very likely to be supplemented by other modes which coexist with it and, in effect, enable it to predominate. Few government bodies, certainly in a democratic structure, have absolute authority. If they have centralized control of

some aspects, this is likely to be accompanied by a policies planning style with indirect control over other aspects of their responsibility. When central government bodies depend on the collaboration of autonomous corporate bodies to implement their policies which are, in turn, dependent on the collaboration of government bodies in meeting their requirements, centralized decision-making and planning and policies planning is likely to coexist with corporate planning. No community is an entity unto itself in modern industrial society. While certain types of decisions can be arrived at at the community level, with resources being allocated accordingly, e.g., for the provision of pre-school or elementary education, others are clearly within the province of a central government or an intermediate level of government, e.g., inter-city highway construction or airport development. In certain circumstances, and in order to obtain more power, community groups may coalesce enabling the emergence of another stronger corporate body which will enter into the negotiation process.

In spite of this obvious convergence we shall analyze each of the planning modes separately with respect to a set of variables which have significant implications for evaluation procedures. By separating out the particular decision-making mode we can better focus on its specific characteristics. We can thus develop an appropriate evaluation strategy, whether it be for a unique or a composite planning mode. The various planning modes will be analyzed in terms of the following variables (See Figure 1):

(1) Degree of centralization of power associated with this planning mode.
(2) Form of control.
(3) Number of clearance points required before implementation (Pressman and Wildavsky, 1973).
(4) Conceptual distance between the decision and the actual intervention in the field.
(5) Emphasis on product or on process.
(6) Number of actions involved.
(7) Accountability i.e., extent to which the citizen can check (obtain information) about what is happening to him/her (latent/ covert or patent/overt planning).
(8) Role of the technical expert.
(9) Who benefits from the mode ?
(10) Assumed consensus in the system.
(11) Breadth of responsibility of the decision-making body.
(12) Opportunity for participation in the decision-making process by interested parties.

We will now discuss these variables in more detail.

Variable/Parameter \ Mode	Command/ Initiatory Planning	Command/ Regulative Planning	Policies Planning	Corporate Planning	Participatory Planning
Degree of Centralization	Centralized	Centralized	Weakly Centralized	Fragmented	Dispersed
Degree of Control	Great Control Budgetary	Great Control Statutory	Indirectly induced -incentives -guidelines -information	Compliance by consent and mutual interest	Voluntary compliance
Number of Clearance Points before Intervention	Few	Can be few or several depending on system hierarchy	At least two stage	Multiple clearance depending on # of corporate bodies	Single or numerous depending on number of participants and issue addressed
Conceptual Distance from Intervention in field	Immediate	Conditional - dependent on initiative of others	Imminent distant depending on degree of consensus	Intercorporate distant or imminent, for corporate bodies immediate	On Issues subject to control - immediate
Emphasis on Product or In process	Product	Product/plan or statutory behavior	Process and product	Process	Process and product
Number of Actors Involved	Fewer than subsequent ones	Fewer than subsequent ones	Potentially numerous	Potentially numerous	Many
Accountability (Transparency)	Patent or latent	Patent or latent	At least partially patent	Patent	Patent
Role of Planner	Technical specialist	Technical specialist	Advisor and simulation analyst	Advocate and negotiator or broker	Advocate and facilitator
Who benefits? (whose interests are served?)	Central decision-making body and "the public"	Central decision-making body and "the public"	Central decision-making body and sectoral/ regional decision-making bodies and "the public"	Corporate bodies and who they represent	Active participants (if they are representative of the entire constituency)
Extent of Consensus in System	Consensus assumed	Consensus assumed	Consensus on basic policies (Not detail)	Consensus on need for negotiated settlement not on issues	Potential for conflict but can facilitate consensus
Breadth of Areas of Responsibility	Responsibility for broad range of areas of public sector activity	Responsibility for broad range of areas of public sector activity	Central body-broad areas of responsibility and sectoral/ regional areas of responsibility of subordinate bodies	Specific areas of responsibility of corporate bodies	Areas of concern of participants as defined by them
Opportunity For Participation by Interested/affected Parties	Limited	Limited	Limited	Greater opportunities for participation by corporate bodies	Based on participation by interested parties

Figure 1. **Variables of a Decision-Making or Planning Process**

1) Degree of centralization of power

The several styles of decision-making and planning have already been defined in terms of the degree of centralization associated with them. We will thus summarize the situation. The command/initiatory and the command/regulative planning modes are by definition centralized. The policies planning mode derives from a weakly centralized decision-making structure. The corporate planning mode occurs when power is decentralized among a small number of corporate bodies. In the participatory mode, power is fragmented and dispersed among many actors.

2) Form of Control

In the initial definitions of the various planning styles, the relevant forms of control have been discussed. In the initiatory version of the centralized command

mode, control is achieved by budget allocations while, in the regulative version, the planning and decision-making institutions can regulate by rule of law. In the case of policies planning control is weaker and indirect. The central decision-making body provides guidelines and decision criteria for the other levels of government, disseminates information to ensure that everyone can operate on the basis of the same data and provides material incentives for those who will follow the guidelines. In the corporate case there is no centralized control, but the behavior of the system is controlled by mutual interest and consent reached through a process of negotiotion and bargining. In the case of participatory planning, control of the system is based on voluntary compliance derived form deliberation by the group and the desire of the participants to continue to be identified with the group on the understanding that, in that way, their interests are best served.

3) Number of clearance points before implementation

By clearance point we mean the points or stations at which decisions have to be made. In hierarchical systems, the clearance points have a vertical spread. When there are many bodies at the same level, each of which would have to be consulted and agree, the clearance points have a horizontal spread. This has significance for determining the number of stages in the evaluation process. In the command initiatory mode of planning there is, by definition, a single clearance point. However, even when budgets are only allocated by an upper governmental body, if application is made for them by a lower level body which decides to made the application, there are in effect two clearance points before intervention. If other bodies have to confirm the application en route there may be more clearance points. An example of this is a program funded at the federal or national level for community development or neighborhood rehabilitation. Allocation of funds is dependent on community or neighborhood initiative with local government confirmation required en route (3 stations).

In the regulative version of the command mode there is, by definition, more than one station. Regulation is established by the central authority in order to impose constraints on the initiatives of other actors. Regulations may be initiated by the central authority or by a lower level authority for which the central authority is the final arbiter. Intervention in the field will, however, usually be dependent on the initiative of public or private bodies, other than the central autority, whose actions are constrained by the regulation. The regulative mode is thus likely to have more clearance points than the initiatory mode. An example of this mode is a higher level planning autority authorized to adopt a statutory land use master plan which is submitted for approval by a lower level authority.

In the case of policies planning, the assumption is that there is a higher level authority responsible for setting guidelines for lower level authorities and providing a common information base and material incentives to achieve compliance. In this case there are at least two clearance points in vertical array although there may very well be more. At the lower level there may be one or many clearance ponts in horizontal array. An example of this is the promulgation of an industrial location or population distribution policy at the national level without sanctions to ensure its implementation. However, the national level can act to encourage implementation by public agencies and private agencies all of whom make their independent decisions but can be influenced by pertinent information and material benefits.

In the corporate planning case, there are as many points of clearance as there are corporate bodies, each of which has to decide whether to adopt the negotiated positions. Examples of this are the decisions of each of the parties involved in ne-

gotiations between corporate bodies concerning a national incomes policy - government ministries, manufacturers associations, trade unions, etc.

In the case of participatory planning, the number of clearance points will depend on the nature of the issue which is being addressed. When the issue under consideration is completely within the province of the participatory group such a decision on neighborhood child-care services by a neighborhood group or a neighborhood clean-up campaign, there is a single clearance point. The situation changes when the activity requires budgets which are not available in the neighborhood and must come from a higher authority - at the city or national level, e.g., for the development of a neighborhood park. In this case there might be several clearance points. A similar situation exists when statutory approval is required form a higher authority, e.g., the need to obtain a building permit for a neighborhood facility form the local government. Another case is one in which the agreement of neighborhood groups in adjoining neighborhoods is required, e.g., for the establishment of a neighborhood service which, because of location or scale, must also serve adjoining neighborhoods. In all the latter cases there are several clearance points.

4) Conceptual distance from decision to intervention in the field

The conceptual distance between the decision and its implementation in the field refers to the imminence of the intervention in the wake of the decision. It affects the time which is likely to elapse between the decision and its implementation and by implication, the degree of certainty that it will be implemented as decided.

In the command-initiatory situation when the decision, once taken, is due to be implemented shortly and by the decision- making institution, the conceptual distance is small and the intervention is imminent, This, however, is not the case for the command-regulative mode, when the regulative decision-making body sets the constraints for action in its regulative role, but does not initiate activity leading to intervention. The conceptual distance, in this case, is dependent on the initiative of other parties and the imminence of intervention may range form immdediate to far off and may never occur.

In the policies planning case, the conceptual distance from the statement of policy guidelines to action in the field reflecting these guidelines could be small. There is only indirect control by the central decision-making body but its effect could be immediate. However, if the incentives to intervene are insufficient to counter conflicting interests of lower level bodies, intervention may be far less imminent and may never come. In the corporate planning situation the responsible decision-making authority of each corporate body can ensure intervention, consistent with its decision, within the orbit of the autonomous activity of the corporation. For those activities which are subject to intercorporate decisions, the imminence of the implementation will depend on the mutual interest of the corporate bodies. If there is somebody interested in the implementation of the decision and agreement has been arrived at with enthusiasm among the negotiating corporate bodies, implementation will be imminent. If this is not the case and the decision is a compromise reflecting a least-bad situation about which nobody is very enthusiastic, and there are no strong pressures from anybody to intervene in the field, intervention may be far from imminent and may never come. In the latter case the conceptual distance between decision and intervention may be very great.

In the case of participatory planning, for those issues whose implementation falls within the control of the participatory group, intervention may be imminent. However, the lead-time to intervention in those activities which require the con-

firmation of another body at a higher level of government or of a parallel participatory institution, may be drawn out causing the conceptual distance between decision and its implementation to be much longer than anticipated.

5) Emphasis on product or process

The question that we are now considering is whether the primary focus of the planning and decision-making style is on the product (i.e., the intervention in the field) or on the process of planning and dicision-making itself. Does this process have inherent value and does it by its nature contribute to change in the field resulting from change in the functioning of the decision-making body and the individuals within it?

In centralized, command-initiatory planning and decision-making the focus is on the product, the intervention in the field resulting from the allocation of the necessary resources. The success of the planning is measured in terms of the substantive intervention in the field which, in turn, leads to the achievement of stated goals. In the command/regulatory mode the emphasis is on the adoption of the regulations which will condition the subsequent intervention in the system thereby allowing that behavior in the field which is consistent with statutory regulations. This presumably will ensure the achievement of the goals of the regulatory plan. Thus the master plan stipulates land uses and permitted densities in order to achieve a certain quality of life, or a particular level of accessibility.

Policies planning is, by definition, oriented to achieving the implementation of particular policies. At the same time it is concerned with setting in motion the apparatus which will lead to their adoption and achievement. Since the decision-making structure is non-coercive, as much thought has to be given to the process of decision-making as its product, including the provision of inducements for the adoption of the policies by other public and private bodies. Since the decision-making system is decentralized, the focus tends to be on policies for a single sector such as education or transportation or housing rather than the comprehensive perspective that frequently characterizes centralized command planning.

Corporate planning, in its inter-corporate dimension, is oriented to compromise between the various corporate entities. It thus tends to be conservative, leading only to incremental change. Central to this mode is its process of decision-making, by negotiation and bargaining. Perhaps more important than the product, which only relates to marginal change, is the maintenance of the corporate system and its way of arriving at decisions. The maintenance and the enhancement of the effectiveness of the negotiating and bargaining procedures is central to corporate planning.

Participatory planning, similarly, places primary emphasis on the participatory process which is highly valued for its own sake and is assumed to contribute greatly to the quality of the human environment for which the planning takes place. The product is also important in this mode but the process is, at least, of equal importance.

6) Number and nature of actors involved in the planning decision-making process

The focus in this section is on the number of parties that are likely to be involved in the decision-making process as well as their homogeneity or variety. This has implications for treatment of the distribution effects in the evaluation process. Is the analysis of the distribution effects of a policy that is under consideration sig-

nificant both from the political point of view as well as from the equity viewpoint? Are planning decisions broadly based with the involvement of a variety of interested parties, including different socio-economic and other interests, or are they made by a small, frequently self-selected, group which tends to be homogeneous in nature?

Command-initiatory planning is usually dominated by a relatively small group of decision-makers who come from higher socio-economic classes and their professional advisors who hail from the same background. The same holds true for command-regulative planning decisions which are usually decided by elected politicians or appointed planning commissioners or upper level civil servants aided by professionals, all of whom tend to belong to the same socio-economic class.

Policies planning, since it is not as centralized, potentially takes in more and a greater variety of actors than the previous modes. Since it involves distinctive sectors of government and different levels of government, it implies that more actors and a greater variety of actors will participate in the decision-making process than in the case of command planning.

In corporate planning, the number of actors will be as varied and as numerous as the number and variety of corporate bodies involved in the process. This can vary from few to many depending on the particular case and this may vary even for similar issues. Thus the location of a potentially polluting industry may be resolved on the basis of negotiations between the owners of the plant and the local authority affected. However, other parties to the negotiation might include health authorities, local, regional and national planning bodies, trade unions (representing potential employees), the local chamber of industry, local citizen groups (representing residents who will be adversely affected), both local and national environmental groups, etc.

In the case of participatory planning, the number of participants in the process is potentially large (depending on the issue) and quite varied, ranging from grass roots interests to high level bureaucrats and planners.

7) Accountability

This refers to the extent that a decision-making body is accountable and responsive to the citizens, expressed first and foremost by the flow of information from the decision-making body **to** the citizens so that the latter can be fully apprised of the plan proposals and their implications. Patent or overt planning ensures a full flow of information to all interested parties, which is not the case with latent or covert planning.

Command planning, of both the initiatory and the regulatory varieties, is frequently carried out in a latent manner avoiding full accountability of the decision-makers and planners to the affected parties. This is the nature of economic planning in the U.S.S.R. but it is also frequently the pattern of land-use master-planning in otherwise democratic societies. These treat the information concerning the plans, while the planning is in process, as being only appropriate for those who belong to the 'in' group on the grounds that thereby land speculation and other ways of taking advantage of the information will be avoided. Alternatively, command planning can be an overt, patent fully accountable process too in which policies are formulated and decided upon in full view of the affected parties, and with their consultation, as has recently occured in the U.K. structure planning process (Alterman, Harris and Hill, 1984).

In the case of policies planning (since it is so decentralized), a broader flow of information concerning prospective policies and their implications is ensured. Conceivably, the information flowing from the central policy-making bodies to the various sectors and hierarchical levels of government can still be treated as privileged information but more of it will inevitably be exposed than in the case of command planning. On the other hand, as with the case of command planning, the decision-making processes of planning institutions could be made completely accountable to their citizens.

In the case of corporate planning, accountability to the general citizenry is nowhere ensured, although the corporate bodies, as such, are usual party to most information flowing from the negotiation process. It can be assumed that each corporate body will share the information that becomes available to it with its members. However, the negotiating process implies that each corporate body does not share all its information with the other corporate bodies. Corporate directors may be accountable to their corporate clients but they are by no means accountable to the clients of other corporate bodies. While there may be some leakage from corporate bodies to the public at large, complete accountability is here wellnigh impossible.
In the participatory mode, participatory groups are fully accountable to their members. Government decision-making processes may not be fully accountable but there are strong pressures to lift the veil from government decision-making and ensure the flow of information to the citizen at large.

8) The role of the professional planner

The role of the professional planner or policy analyst must perforce vary in accordance with the planning mode. In the command mode, both initiatory and regulative, the planner is very much the technical specialist. He is primarily engaged in information collection and processing and thereby trying to understand the effect of the various courses of action under consideration on the total system, its goals and its behaviour. There is little emphasis on interpersonal skills and activity, particularly if public participation is not mandated.

In the policy mode the professional planner is cast as a policy advisor. Here the responsibility is not to collect information about the entire system but to identify key policy variables. By pursuing these policies the system can be directed in the desired direction. The advisor must assemble information about these key variables and simulate the effects of courses of action which involve them. The advisors need interpersonal skills as well as analytical ones, in order to fulfil their duties adequately.

In the case of corporate planning one can distinguish between the role of planners working for the corporate bodies and those in the central planning office at the intercorporate level. Corporate planners are essentially advocates for the interests of the corporations with whom they are identified. They need analytical skills to analyze the effects of particular negotiating positions and outcomes on their client corporate body. At the central planning office, at the intercorporate level, the planner has to fulfil the role of broker in the negotiating and bargaining process. Planners require inter-personal skills for this purpose as well as technical skills to enable them to draw up draft documents for decisions backed by technical analyses and forecasts.

In participatory planning the planner must be a facilitator of the participatory groups, enabling the community groups to plan for themselves by providing the necessary information and professional guidance while not dominating the process.

Where necessary the planner must mediate among competing groups. The process may not require sophisticated methods of analysis for provision of information. However, the planner requires skills in community organization and interpersonal communication.

9) Principal beneficiaries of the planning mode

The question here is whose interests are primarily served by a particular mode of decision-making and planning ? Who are the main beneficiaries of this mode as a result of control over the decision-making process, the role that the decision-makers attribute to themselves and the perspective of the public interest that is assumed ?
The command mode, initiatory or regulative, serves the interest of the central decision-making body who attribute to themselves (as a result of legitimate democratic procedures or otherwise) the responsibility for expressing and representing the public interest. They enjoy the decision-making power which derives from this responsibility and, at the same time, serve the broader public at large. This latter contention is open to debate because it is questionable whether a broad public interest indeed exists and, if it exists, whether a particular group of people represent it. If this claim is substantiated in fact as well as symbolically, then the interests of the broader public (if such exists) can be assumed to benefit from this mode.

In the policy planning mode, the decison-making power is spread between the central decision-making bodies and the sectoral and regional decision-making bodies. The benefits to be derived from the direct exercise of the power of decision are therefore spread among a larger group of people. Because policies planning has no coercive means, it may be assumed that compliance with it reflects a convergence of interests between the central authority and the sectoral and regional bodies. As in the case of the command planning mode, the decision-making bodies purport to represent the broad public interest of their constituency, however defined.

The corporate planning decision-making process clearly serves the interest of the corporate bodies and those people that they represent and it is they who benefit from this mode of planning. Even at the intercorporate level, where the planners are engaged in trying to mediate among the corporate bodies to arrive at an acceptable solution, they are serving the agregate of the corporate bodies and those that they represent rather than the more nebulous concept of the public at large.

The participatory mode of decision-making and planning ostensibly serves the interests (and hence benefits) each of the community participatory groups involved in the process. However, although these groups might purport to represent the entire constituency, the planning and decision-making process represents, first and foremost, the active participants in that constituency.

10) The extent of consensus that is assumed to exist in the system

The assumed existence of a consensus about what is in the public interest is a key variable for the development of an evaluation strategy. In the command planning mode, both the regulatory and the initiatory cases, a consensus on the public interest is assumed to exist with the central decision-making body being responsible for articulating the consensus, both with respect to the general policy objectives and with respect to the detailed policies.

In the policies planning case, there is assumed to be a consensus on basic policy

issues which is articulated by the central decision-making body. However, there may not be a consensus on the implications for detailed policies. Differences of opinion may arise about detailed policies when these are approached from the point of view of different sectors or regions. Thus, all may agree on the need to stem center city congestion but the highway planning department may see its solution in the development of a highway system while the city planning department might foster a rapid transit system. All might agree on a policy for urgently developing new energy resources. In the one region they might push for the development of the coal resources which are abundant in that region, thereby also providing additional employment for the depressed region. In another region, blessed with abundant water resources, the policy might emphasize putting the investments into the development of a hydro-electric system.

In the corporate planning mode there may be no consensus among the various corporate bodies on basic issues. However, there is a consensus on the need for a negotiated settlement among the corporate bodies. In the participatory planning mode there is potential for conflict within participatory bodies but this can frequently be judiciously avoided by consensus-promoting activities and discussion stimulated by the planners.

11) The breadth of areas of responsibility of the decision-making body

Centralized command-initiatory planning and decision-making bodies usually take upon themselves a broad range of areas of responsibility. Being responsible for deciding about budgetary allocations they have to determine priorities among the many competing sectors. In making such decisions, they have to take into consideration interdependence and interactions among the sectors. Command-regulative planning bodies usually base their proposals on a comprehensive analysis which considers priorities in and between the various sectors of activity as well as the interactions between them. Here, too, comprehensive responsibility implies comprehensive planning.

The centralized policy planning body assumes responsibility for the broad sweep of geographic areas and sectors of activity, as in the centralized command case. However, in this case, the central policy-making mode is complemented by sectoral and regional planning bodies who assume responsibility for deciding on policies for these sectors and regions respectively.

In the corporate planning mode, each corporate body assumes responsibility for its specific area of concern, there being no central body which has broad responsibility as in the previous modes. In participatory planning, each participatory group assumes responsibility for those activities which are defined by the participatory group as being in their province of concern.

12) The opportunity for participation in decision-making bodies

Centralized command planning is not noted for facilitating participation of interested and affected parties. In some cases, however, as in U.K. structure planning, such participation is mandated, but this tends to be exceptional. Command planning, with power vested in the central decision-making body, is by definition not oriented to facilitate public participation.

Policies planning is similarly not based on the assumption of significant public participation. Policies planning may be more decentralized than command planning, but decentralization to the sectoral and regional levels does not ensure participation.

Corporate planning and decision-making tends to provide opportunities for involvement by members of corporate bodies appointed and elected to the task from within the corporate body.

Participatory planning, by definition, provides considerable opportunity for participation at the grass roots. In fact, the very raison d'etre of this mode is its participatory component which conditions the entire orientation of the mode.

With this we complete our discussion on the variables which characterize the planning and decision-making modes and will now consider the implications of these characteristics of the decision-making process for developing the evaluation strategy. In suggesting that the evaluation strategy should reflect the nature of the decision-making process, one must not forget that other factors may be just as important as the decision-making characteristics in determining this strategy and thus need to be taken into consideration as well. Among such factors are the perceived urgency of the problem that has to be dealt with; the knowledge that exists about the system that is being planned and hence the confidence in the forecasts about the expected results of future policies; the scale of the projects or the policies under consideration, in itself and as part of the total budgetary allocation. We will not address these aspects at this point but we mention them as aspects which should not be neglected when deciding on evaluation procedures.

RELATIONSHIP BETWEEN DECISION-MAKING CONTEXT VARIABLES AND EVALUATION VARIABLES

We will now review each of the decision-making variables which varied by planning decision-making mode and see what they might imply for the evaluation strategy (See Figure 2).

The following relations are distinguished:

1) Relationship between degree of centralization of power and perception of the public interest

The perception of the public interest is an important consideration in deciding what emphasis is to be placed in the evaluation on the identification of multiple interest groups and how they will be affected by the policies under consideration. The highly centralized command mode implies a unitary perception of the public interest. This perception assumes the existence of a public interest which supercedes particularistic interests for those responsible for making decisions in the

1) Relationship between degree of centralization of power and perception of the public interest	
Highly Centralized	Single Central Public Interest
Weakly Centralized	Multiple Public Interest with Central Dominant
Diffuse	Multiple Public Interest

2) Relationship between degree of control and degree of certainty in assessment	
Great Control (budgetary or statutory)	Relative Certainty
Indirect Control	
Control of Mutual Interest	
Voluntary Compliance	Relative Uncertainty

3) Number of clearance points before implementation and number of stages in the evaluation	
Single Clearance Point	Single-Stage Evaluation
Few Clearance Points	
Many Clearance Points	Multi-Stage Evaluation

4) Conceptual distance between decision-making and intervention in the field and degree of certainty in assessment	
Immediate Intervention	Relative Certainty
Distant Intervention	Relative Uncertainty

5) Elapsed time between decision-making and intervention in field and expression of time preference	
Immediate Intervention	No Need for Assessing Time Preference
Distant Intervention	Time Preference to Be Accounted For

6) Emphasis on product or on process and stress on ex-ante (one time) or continuous or ex-post evaluation (3)	
Product	Ex-ante (and ex-post)
Product and Process	Ex-ante and Continuous
Process	Continuous Evaluation

7) Number and nature of actors involved and degree of emphasis on equity or distribution effects	
Single Actor	No Emphasis on Distribution
Multiple Actors	Great Emphasis on Distribution

8) Accountability and the relative importance of comprehensibility of evaluation to all actors (4)	
Accountability Patent	Methodology Must Be Clearly Comprehensible to All Actors
Latent	Methodology Can Remain In Domain of Expert

9) Role of the planner-technical expert and level of sophistication of evaluation methodology (5)	
Technical Specialist	Sophisticated Technical Methodology
Technical Expert and Advocate	
Negotiator	
Facilitator (organizer)	Simple Methodology

10) The effects on those whose interests are served should be evident in the evaluation procedure	
	Unitary Public Interest
	Central Decision-Making Body
In Accordance with Mode of Decision Making	Sectoral/Regional Decision-Making Body Corporate Bodies
	Participatory Public
	All the Publics

11) Extent of consensus in system and: a) Possibility of unified objective function, b) Feasibility of an optimizing or satisficing solution		
Complete Consensus Assumed	Emphasis on Unified Objective Function	Optimal Solution
Complete Absence of Consensus	Impossibility of Achieving Unified Objective Function	Satisficing Solution

12) Areas of responsibility of decision-making body and extent of comprehensibility of evaluation	
Broad Responsibilities	Comprehensive Evaluation
Sector/Regional Responsibilities	
Areas of Responsibility of Corporate Bodies	Narrow Sectoral Evaluation Narrow Interest Evaluation

13) Extent of institutionalized public participation and interactive nature of the evaluation process	
Little Peer Participation	Centralized Evaluation (Non-interactive)
Great Deal of Peer Participation	Interactive Evaluation Techniques

Figure 2. **Implications of characteristics (variables) of the decision-making process for evaluation variables**

public sector. Thus, in centralized planning, the major concern is how 'the' public interest is affected. This is in fact, the basis for application of cost-benefit analysis methodologies. However, as Lichfield has shown in his analysis of the Third London Airport decision (Lichfield, 1971), the Royal Commission which can, in this case, be considered as an archetypal central planning institution, would have done better had it ad;opted a multiple interest analysis instead of a methodology which assumed a unitary public interest. In those decision-making modes which assume a diffusion of power, there is no single central public interest but the public interest is assumed to be composed of a composite of the interests of multiple publics. In such cases, the point of departure of the evaluation process could very well be the identification of the multiple publics and an assessment of how their interest will be affected. When the decision-making process is weakly centralized, one can argue for balance between an assessment of the effects on what is assumed to be the central public interest and the effects on multiple particularistic interests.

2) Relationship between degree of control and degree of certainty assumed in policy assessment

One of the key considerations in evaluation strategies is the degree of certainty that can be attributed to the implementation of the policy. When there is a high level of control there is relative certainty that the policy will be implemented. As the degree of control decreases, as is the case when it is indirect or is dependent on the convergence of mutual interests, there is obviously less certainty about the extent of implementation. The highest level of uncertainty exists when control is achieved by voluntary compliance. As we have previously noted the various planning decision-making modes are characterized by varying degrees of control over policy implementation ranging from a high degree of control over budgetary allocation or statutory regulation to very low levels of control by voluntary compliance.

Uncertainty (when it is impossible to attibute a probability level) can be considered in the evaluation strategy by methods derived from game theory; risk (when probability levels of uncertainty can be assigned) can be treated by methods derived from probability theory.

Obviously, the uncertainty about the implementation of future policies and their impacts is not only dependent on the degree of control and on whether the policies will be implemented as formulated. The level of certainty is also dependent on the availability of adequate theory and empirical knowledge. For instance, the effects of policies in sectors related to the natural environment or physical systems such as water resources systems or transportation systems (for which there are well established scientific theories) can be predicted with a high degree of confidence. By contrast, the expected effects of social policies can only be predicted with a much lower level of certainty since the theoretical understanding is much less developed.

Other elements which may affect the level of outcomes and their assessment include the stability in time of policy objectives and the stability of social systems and existing technologies. In any case, the above types of uncertainty have to be

considered together with uncertainty arising from the implementation of policy.

3) The number of clearance points before implementation and the number and nature of stages in the evaluation study

As previously defined, clearance points refer to points of decision which policies have to pass through before they are adopted. As we have shown, the number of such clearance points can vary from a single clearance point to numerous. At each clearance point, the perspective of the decision environment might change. Thus the initial decision-making body required to confirm a policy may have sectoral responsibility. The next body that has to confirm the policy may be at the local or regional level but with general responsibility for policies adopted in all sectors. The next clearance point might be at the national level but, once more, with a sectoral responsibility. Finally, the policy may have to be confirmed at the national level by a body responsible for comprehensive policy-making such as the National Budgetary Department of the Ministry of Finance or the National Planning Department.

At each clearance point, the decision may be approached differently with a different set of factors being taken into consideration depending on the mandate of the decision-making body and its defined responsibility. At each level, a different type of evaluation study may thus be required. A lower-level sectoral decision followed by a higher level decision by a decision-making body with comprehensive responsibility may require first a sectoral functional and cost-effectiveness analysis followed by second-stage comprehensive multi-objective evaluation. The number of stages in the evaluation study should thus reflect the number of clearance points in the decision-making process and the nature of the evaluation studies at each stage should reflect the perspective and nature of the responsibilities of the decision-making bodies.

4) The conceptual distance between decision-making and intervention in the field and the degree of certainty in assessment

The degree of certainty that a policy will be implemented as adopted is influenced by the factors enumerated above. In addition, we suggest that it will also be affected by the conceptual distance and hence the time that is likely to elapse between the decision to adopt the policy and its implementation in the field. As we have demonstrated, in a decision-making mode like the command-initiatory type, the conceptual distance, and hence elapsed time, to intervention could be small and hence implementation will be more certain. In the command-regulative case where intervention is dependent on the initiative of others, the elapsed time to intervention is likely to be longer and hence the nature of the intervention is less certain. In the case of policies planning, corporate planning and participatory planning, the conceptual distance between decision-making and policy implementation in the field can be small or large as noted above. These have direct consequences for the degree of certainty of implementation.

5) The elapsed time between decision-making and intervention in the field and assessment of time-preference

This point may be common-place but we shall state it nevertheless since we are reviewing the various variables in evaluation studies. If the benefits and costs which are to accrue as a result of the intervention are expected to occur at different times in the future then it is necessary to account for this in the evaluation study by analysis of time preference. This is usually done by applying a discount rate to future costs and benefits and discounting them to present value. This

traditional treatment of costs and benefits may have an additional component depending on the mode of decision-making. The conceptual distance between the adoption of the policy and intervention in the field and hence the time which is expected to elapse between decision and intervention, should be considered as an additional factor to be taken into consideration when assessing time preference and hence, present value. As we have pointed out in our previous discussions this elapsed time is expected to vary depending on the mode of decision-making.

6) Emphasis on product or on process and stress on ex-ante, continuous or ex-post evaluation

By ex-ante evaluation we mean the evaluation of alternative courses of action prior to intervention in the field in order to inform the decision-making process. By ex-post evaluation we mean the evaluation of the policy or the plan after it has been implemented in the field by addressing questions of what has been implemented in the field; what considerations influenced the policy decisions; how they were taken; and what the impact in the field were? By continuing evaluation we mean the incorporation of evaluation into a continuing planning and decision-making process. The decision-making process and implementation in the field are monitored and impacts in the field are periodically assessed in order to inform the ongoing planning and decision-making.

We have demonstrated that in all of the planning modes there is differential concern with product and process, depending on the mode of decision-making. Where the emphasis is on product there is a primary need for the ex-ante evaluation of alternative pruducts in order to aid the decision-making process by means of cost-benefit analysis, cost-effectiveness analysis or any of several multiple objective methods of evaluation. After the product is delivered there might be a recognized need for ex-post evaluation which emphasizes whether the product or outcome was delivered as planned and what the impact of the product was.

Where the emphasis is on the planning and decision-making process, as such, it can be evaluated in terms of its efficiency, its effectiveness, its accountability, its representativeness and other criteria. This type of assessment is complementary to the evaluation of product and could be a component of ex-post or continuous evaluation procedures (Alterman, Carmon and Hill, 1984).

Thus, in the case of initiatory demand planning, where the emphasis is on product, the relevant types of evaluation are ex-ante evaluation and, if the need is recognized, ex-post evaluation. All the other modes of planning are also concerned with product but in policies planning, corporate planning and participatory planning, there is considerable interest and urgent need for evaluation of the planning and decision-making process itself (Alterman, 1983; Palumbo and Harder, 1981).

The approach to evaluation might be different depending on whether the plan is for a physical facility or for a social service or an institutional change. In the case of the physical plan, particularly if it pertains to the construction of a facility in a limited period of time, ex-ante evaluation will suffice, possibly followed by ex-post evaluation. However, if the product is itself in a constant state of change, as is the case for a social institution or a social service, there is a need for continuous evaluation which assesses the changes which are taking place.

7) Number and nature of actors involved and the degree of emphasis on equity or distribution effects

Policy goals frequently call for an improvement in the equity of the distribution of

goods and services or at least that equity distribution effects be taken into consideration in the evaluation process. A second reason for tracing distribution effects may be a political one. The likelihood of the acceptability of a particular course of action might be dependent on taking into consideration the interests of all the significant actors in the planning and the decision-making process. If there is a single target group or if the actors are largely homogeneous in terms of their relevant socio-economic and other characteristics then the distribution effects may be of relatively little significance from the political point of view. However, if there are multiple and heterogeneous groups of actors then it may be necessary to take the distribution of the planned goods and services into consideration in order to make the plan politically feasible. In practice only in command planning can such political considerations be neglected because of the homogeneity of the decision-making body. However, even here the relatively limited heterogeneity of the decision-making body may require treatment of distribution effects in order to ensure tant the plan will be politically acceptable. This will certainly be the case for all the other modes of planning and decision-making - policies planning, corporate planning and participatory planning as we have previously demonstrated.

8) Accountability and the comprehensibility of the evaluation to all actors

We have defined accountability as being related to the flow of information from the decision-making body to the citizen so that the citizen may be fully acquainted with policy proposals and their implications. Patent or overt planning ensures a free flow of information from planner to the public while latent or covert planning discourages a free flow of information.
In the case of patent planning (which as has been pointed out could exist for all of the planning and decision-making modes), it is important that the evaluation methodology be clearly comprehensible to all actors in the decision-making process including the public at large. This means that a very sophisticated methodology, which requires expert technical knowledge (to follow and to understand its outcomes) is not acceptable.

In the case of latent or covert planning, with a low level of accountability to the public, the intricacies of the methodology can remain in the domain of the expert since it does not have to be communicated to the public at large. As we have pointed out, such latent planning is particularly evident in the case of elitist command planning. But the decision-maker would want to know its minute details, perhaps more so than in patent planning.

Accountability is particularly needed where the planning involves some highly controversial and publicly visible issues, e.g., the case of the Third London Airport. The Royal Commission for the Third London Airport employed cost-benefit analysis to compare alternative sites (Flowerdew, 1972). This required their making some uncommon assumptions in order to arrive at measures of the economic value of items such as aircraft noise, travel in leisure time etc. There was severe criticism of the method and its resultant conclusions (Mishan, 1970; Self, 1970). Thus, whereas in other circumstances, under conditions of command planning, the planning could be latent/covert, where the issues are controversial and in the public eye, the methodology has to be clear and uncontroversial.

9) Role of the planning/technical expert and level of sophistication of the evaluation methodology

As we have demonstrated, the role of the professional varies according to the planning mode. In the command mode, the planner is the technical specialist. In

the policy planning mode, the professional is cast as a policy advisor requiring technical skills as well as interpersonal ones. In the corporate mode, the professional working for corporate bodies operates as an advocate planner and planning analyst while in the intercorporate case the professional operates as a broker or a negotiator. In the participatory mode, the professional works as a facilitator.

When the planning process is primarily the responsibility of a technical specialist or analyst, there is room for a relatively sophisticated technical methodology of evaluation. However, when the professional is required to work closely with the public as mediator or facilitator the methodology of evaluation must perforce be simple and comprehensible to the public.

Another consideration is that the more controversial the issue, the simpler the preferred methodology should be, so that its implications can be clearly understood and related to those involved.

10) The effects on those whose interests are served should be evident in the evaluation procedure

As we have indicated the various planning modes ostensibly also serve the interests of those who are involved in the decision-making process. Since these people and institutions exercise power over the decision-making process, it makes sense to include in the assessment an analysis of the effects on the interests of those who control the decision-making process. To a social scientist this might appear to be a cynical approach to evaluation. However, it is proposed for inclusion since in any event those in control will take their own interests into consideration in their decisions. Explicitly expresssing these in the evaluation makes them part of the total set of information on which the evaluation is based.

Thus, in the command decision-making mode, the interests of the central decision-making body should be identified; in the policies planning mode, the interests of the sectoral/regional decision-making bodies as well as those of the central decision-making body should be identified. In the case of the corporate planning mode, the effect of the policies under consideration on the interests of corporate bodies should be identified. In the participatory mode of planning and decision-making, the interests of the community participatory groups must be traced.

As evidence for the importance of tracing the effects on the interests of the decision-making body in the evaluation procedure, let us once more cite the case of the Third London Airport Study. In this case, the aggregate cost-benefit study did not trace the distribution of effects on affected parties. In particular, it did not take into consideration the fact that the interests or large landowners in the area of Cublington, which was the preferred site, would be adversely affected and its possible implications for the implementation of the plan. The influence of these landowners on the Conservative government, the final arbiter, eventually contributed to the recommendations of the Royal Commission, which were based on the cost-benefit analysis, being overturned. In addition, the cost-benefit analysis dit not take into consideration the destruction of historic areas in Cublington resulting from the proposed airport location. Supporters of historic preservation who had the government's ear were instrumental in influencing the government to reject the original proposals.

11) Extent of consensus in the system and:
(a) Possibility of obtaining a unified objective function
(b) Feasibility of applying an optimizing or satisficing solution

Many evaluation studies assume the existence of a consensus about objectives and

hence that a unified objective function exists. In many cases this is subject to question. The degree of consensus existing among the actors who are party to a decision determines whether a unified objective function can be assumed.

As we have demonstrated, the various planning mode imply a varying amount of consensus. In the command planning case a consensus, based on a unitary public interest, is assumed and hence a unified objective function can be assumed. In the policies planning case there may be a consensus on basic policy issues but not about how these are applicable at the field level since there is an assumed separation of powers and hence, possible conflict of interest between sectors and regions. However, in planning for particular regions or sectors but not across them, it is possible to arrive at a unified objective function. In the corporate mode, as we have pointed out while there may be consensus on the need for negotiating among the various corporate bodies, there is no consensus about common interests, and thus a unified objective function is not feasible. In the participatory mode, consensus may or may not exist.

When there is a unified objective function, it is conceptually feasible to arrive at an optimal solution based on the agreed set of objectives. However, when there is no consensus, an optimal solution is not conceptually feasible and the most suitable solution is that one in which all the interested and affected parties see themselves at least adversely affected, but which is unlikely to be the optimal solution from the point of view of any of their particular interests. This represents a satisficing solution rather than an optimizing one (Werczberger, 1983; Lomovasky and Hill, 1984).

12) Area of responsibility of the decision-making body and the extent of comprehensiveness of evaluation

We have demonstrated that the breadth of responsibility of the decision-making body will vary with the planning mode. The areas of responsibility vary from a perceived broad range of responsibility in the centralized command mode to sectoral or regional responsibility in the policies mode. Corporate bodies have narrower responsibilities while in the participatory mode the responsibilities are also more narrowly defined.

The centralized command planning and decision-making bodies are obliged to attempt a comprehensive evaluation reflecting the breadth of their responsibility while the decision-making bodies with sectoral or corporate responsibilities will find evaluation with a narrower perspective adequate for their needs.

13) The extent of institutionalized public participation and the interactive nature of the evaluation process

The opportunity for public participation in the planning and decision-making process varies with the planning mode. Whereas the command mode does not inherently require public participation, participatory planning is by definition so oriented.

In the participatory mode and others which facilitate participation, there is a need to change the thrust of evaluation methods which have been developed for largely non-participatory modes and which are therefore inherently non-participatory (Sager, 1981). Public participation calls for interactive evaluation techniques which can provide an input into the evaluation process of participatory groups (Lomovasky and Hill, 1984). In this way the evaluation can incorporate the subjective perceptions of the individuals who are affected including their perception of

their problems and policy alternatives, their goal preferences and preferences with respect to policy solutions. The non-participatory central command bodies do not provide an opportunity for interactive planning.

CONCLUSION: TOWARDS MORE CONTEXT RESPONSIVE EVALUATION METHODS

The ultimate test of policy evaluation is its usefulness in helping decision-makers arrive at more rational and relevant decisions (Hill, 1984) - decisions that are implementable and that will ultimately be implemented.

As Barras and Broadbent (1981) have shown many of the evaluation studies which were mandated by the Department of Environment for the preparation of structure plans in the U.K. during the 1960's and 1970's failed in their purpose. They did not contribute to a rational decision-making context and did not interact sufficiently with the political process.

Evaluators should be cognizant of the decision-making context and design the evaluation methodology accordingly. This means that the approach to the evaluation problem must be sufficiently flexible so that the appropriate evaluation methodology can be developed and employed. It also means that evaluators may have to add a measure of political sensitivity to their technical expertise.

If it does not sufficiently inform the decision-makers and the public so that they can use the information provided in order to arrive at more rational decisions, evaluation is an academic exercise. For this purpose, evaluation will have to be more context responsive.

NOTES

(1) This multiple interest perception of the public interest in evaluation is also evident in some papers by Davos, Smith and Nienberg (1979), relating to power plant siting; by Sobral, Hipel and Farquhar (1981) relating to solid waste management; and by Werczberger (1983).

(2) The reference is to papers by P. Nijkamp and P. Rietveld entitled "Hierarchical Multi-objective Models in a Spatial System" pp. 163-186; by Y.Y. Haimes and K. Tarvainen entitled "Hierarchical-Multiobjective Framework for Large Scale Systems" pp. 20-234; by P.L. Yu and L. Seiford entitled "Multistage Decision Problems with Multiple Criteria" pp. 235-244.

(3) Could depend on whether physical plan or plan for social services or institutional change.

(4) Particular emphasis on accountability for highly controversial and visible issues.

(5) When the more controversial the issue is, a simpler methodology may be preferred.

REFERENCES

Alterman, R., Implementation Analysis in Urban and Regional Planning: Toward a Research Agenda, In: P. Healey, G. McDougall and M. Thomas (eds.), **Planning Theory: Prospects for the 1980's**, Oxford, Pergamon Press (1982) 225-245.

Alterman, R., Recent Literature on Implementation Analysis: the Contours of an Emerging Debate, **Journal of Planning Education and Research**, 3, 1 (1983) 63-65.

Alterman, R., N. Carmon and M. Hill, Integrated Evaluation: A Synthesis of Approaches to Broad-Maim Programs, **Socio-Economic Planning Sciences** (1985) (forthcoming).

Alterman, R., D. Harris and M. Hill, The Impact of Public Participation on Planning: The Case of the Derbyshire Structure Plan. **Town Planning Review**, 55, 2 (1984) 177-196

Barras, R. and Broadbent T.A., A Review of Operational Methods in Structure Planning, In: D. Diamond and J.N. McLoughlin (eds.). **Progress in Planning**, Oxford, Pergamon Press, 17, 2/2 (1982) 57-238.

Cain, G.C., and R.G. Hollister, Evaluating Social Action Programs, In: H. Haveman and J. Margolis (eds.), **Public Expenditure and Policy Analysis** (2nd Edition), Chicago, Rand McNally, (1977) 259-281.

Cochrane, J.L., and M. Zeleny (eds.), **Multiple Criteria Decision Making**, Columbia, University of South Carolina Press (1973)

Dagupta, A.K., and D.W. Pearce, **Cost-Benefit Analysis, Theory and Practice**, London, McMillan (1972)

Davos, C.A., C.J. Smith and M.W. Nienberg, An Application of the Priority-Trade-Off-Scanning Approach, Electric Power Plant Siting and Technology Evaluation, **Journal of Environmental Management**, (1979) 105-125.

Flowerdew, A.D.J., Choosing a Site for the Third London Airport, In: R. Layard (ed.), **Cost-Benefit Analysis**, Harmondsworth, Middlesex, Penguin Books (1972) 431-451.

Friedmann, J., **Retracking America**, Carden City, New York, Anchor Press/Doubleday (1973)

Goldman, T.A. (ed.), **Cost-Effectiveness Analysis**, New York, Praeger (1967)

Hill, M., A Goals-Achievement Matrix for Evaluating Alternative Plans, **Journal of the American Institute of Planners**, 34, 1 (1968) 19-28.

Hill, M., **Planning for Multiple Objectives**, Monograph Series No. 5, Philadelphia, Regional Science Research Institute (1973)

Hill, M., Can Multiple Objective Evaluation Enhance Rationality in Planning in M. Breheny and A. Hooper (eds.). **Rationality in Planning**, London, Pion (1985) (forthcoming)

Koslowski, J. and T.J. Hughes, **Threshold Analysis, A Quantitative Planning Method**, London, Architectural Press (1972)

Levin, H., Cost-Effectiveness in Evaluation Research, In: M. Guttentag and E.L. Struening (eds.), **Handbook of Evaluation Research**, No. 2. Beverley Hills, Ca., Sage Publications (1975) 89-124.

Lichfield, N., Cost-Benefit Analysis in Town Planning, A Case Study - Swanley, **Urban Studies**, 3, 3 (1966) 215-249.

Lichfield, N. (1971), Cost-Benefit Analysis in Planning, A Critique of the Roskill Commission, **Regional Studies**, 5, 3 (1971) 157-183.

Lichfield, N., P. Kettle and M. Whitbread, **Evaluation in the Planning Process**, Oxford, Pergamon (1975)

Lomovasky, G., and M. Hill, The Minimal Requirements Approach to Policy Modelling in Public Planning, **The Journal of the Operational Research Society** (1985) (forthcoming).

Major, D., Multi Objective Water Resources Planning, Water Resources Monograph No. 4, Washington D.C. American Geophysical Union (1977)

Marglin, S.A., **Public Investment Criteria**, London, George Allen and Unwin (1967)

Meyerson, M. and E.C. Banfield, **Politics, Planning and the Public Interest**, New York, The Free Press (1975)

Miller, D.H., Equity and Efficiency Effects of Investment Decisions and Multi-criteria Methods for Assessing Distributional Implications. (this volume)

Mishan, E.J., **Cost-Benefit Analysis**, London, George Allen and Unwin (1971)

Mishan, E.J., What is wrong with Roskill ?, **Journal of Transport Economics and Policy**, 4, 3 (1970) 221-234.

Nijkamp, P. and J. Spronk (eds.), **Multiple Criteria Analysis**, Aldershot, Hampshire, Gower (1981)

Nijkamp, P. and A. Van Delft, **Multi-Criteria Analysis and Regional Decision Making**, Leiden, Martinus Nijhoff (1977)

Nijkamp, P. and J.B. Vos, A Multicriteria Analysis for Water Resource and Land Use Development, **Water Resource Research**, 13, 3 (1977) 513-518

Palumbo, D.J. and M.A. Harden (eds.), **Implementing Public Policy**, Lexington, Mass, Lexington Books. Introduction (1981) ix-xvi

Peskin, H.M. and E.P. Seskin, **Cost-Benefit Analysis and Water Pollution Policy**, Washington, The Urban Institute (1975)
Poulton, M.C., A Land Use Evaluation Technique for Decision Makers, **Regional Studies**, 6, 2 (1981) 85-96.

Pressman, J.L. and A. Windavsky, **Implementation - How Great Expectations in Washington are Dashed in Oakland, etc.**, Berkeley, Ca, University of California Press (1973)

Rossi, P.H. and H.W. Freeman, **Evaluation, A Systematic Approach** (2nd edition),

Beverley Hills Sage Publications (1982)

Schamberg, N.P., The City of Simi Valley, California Net/Benefit Assessment Process, Washington D.C., The American Institute of Planners (1977)

Schermer, J.H., Interest Group Impact Assessment in Transportation Planning, **Traffic Quarterly**, 29,1 (1975) 29-49.

Self, P., Nonsense on Stilts - The Futility of Roskill, **New Society**, 2 July (1970) 8-11.

Sengupta, S.S., M.L. Podrebarac and T.D.H. Fernando, Probabilities of Optima in Multi-objective Linear Programs in J.L. Cochrane and M. Zeleny (eds.), **Multiple Criteria Decision Making**, Columbia University of South Carolina Press (1973) 217-235.

Sobral, M.M., K.W. Hipel and G.J. Farquhar, A Multi-Criteria Model for Solid Waste Management, **Journal of Environmental Management**, 12 (1981) 97-110.

Spronk, J., Capital Budgetting and Financial Planning with Multiple Goals, In: P. Nijkamp and J. Spronk (eds.), **Multiple Criteria Analysis**, Harmondsworth, Gower Publishing (1981) 25-36.

Steiner, P.O., The Public Sector and the Public Interest, in R.H. Haverman and J. Nargolis (eds.), **Public Expenditure and Policy Analysis**, Chicago, Rand McNally, (1977) 27-66.

Stuart, D.G., Interactive Sensitivity Analysis in Transportation Plan Evaluation, **Transportation** 3,3 (1972) 289-310.

Voogd, H., Stochastic Multicriteria Evaluation by Means of Geometric Scaling, **Delft Progress Report**, 5 (1980) 112-125.

Weisbrod, B.A., Collective Action and the Distribution of Income, A Conceptual Approach in R.H. Haverman and J. Margolis (eds.), **Public Expenditure and Policy Analysis**, 2nd edition, Chicago, Rand McNally (1977) 105-132.

Werczberger, E., Multiperson Multitarget Decision Making, Using the Versatility Criterion, **Regional Science and Urban Economics**, 13 (1973) 103-113.

U.K. Dept. of Transport, Trunk Road Proposals - A Comprehensive Framework for Appraisal, The Standing Advisory Committee on Trunk Road Assessment. Chairman, Sir George Leitch, London, Her Majesty's Stationery Office (1979)

Zeleny, M. (ed.), **Multiple Criteria Decision Making**, Berlin, Springer Verlag (1975)

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EQUITY AND EFFICIENCY EFFECTS OF INVESTMENT DECISIONS: MULTICRITERIA METHODS FOR ASSESSING DISTRIBUTIONAL IMPLICATIONS

D.H. Miller
University of Washington, Seattle
Dept. of Urban Planning

INTRODUCTION

A major source of the complexity facing public decision making is the increasing insistance that both the efficiency effects and the social redistribution effects of these decisions be taken into account. Public intervention in the market economy is usually justified by the existence of market failures of various sorts, and this intervention takes the form of public supply of goods and services and of regulating behavior in the private sector. Socially acceptable accounting for the equity effects of allocational decisions - that is, assessment of who benefits and who bears the costs - has not been a function of the market economy, and thus in itself is one form of market failure.

All public investments, and controls, generate both efficiency effects and redistribution effects. Traditionally, efficiency effects have been the focus of planning evaluation and policy analysis. For example, site acquisition and site preparation costs for alternative facility locations regularly are assessed for their **internal** efficiency effects. Over the last few decades, we have begun to include the **external** efficiency effects of investments in our analysis. These external effects include second round economies such as reduced transport costs to industry resulting from highway and rail improvements, and spill-overs such as likely noise pollution impacts on activities adjacent to new highway facilities. These external effects are now conventionally included in benefit-cost and other forms of evaluations.

While we have made technical progress in assessing the relative workability of planning alternatives, we have also become increasingly aware that every public decision results in a new distribution of costs and benefits among the people making up the population of the affected area. As Mel Webber (1969: 286) points out, "These redistributive consequences, commonly external to the subsystem being planned, affect various non-client groups, each in different ways and to different degrees". Webber attributes our current attention to "community values" to our recognition of social pluralism, and to growing public debate over the equity effects of public actions. As will be discussed later, we have much yet to accomplish in developing and applying adequate means for assessing the redistributional effects of plans and policies.

APPROACHES TO TREATING BOTH EQUITY AND EFFICIENCY

Attempting to address both efficiency and equity criteria in complex decision making is a rather heroic undertaking, and certainly not a topic to be exhausted in a few pages. This is because efficiency and equity have tended to be the "oil and water" of economics: they do not mix well in most operational contexts. While welfare economics includes a number of overarching objectives - including desirable rate of growth, full utilization of resources, equity, and allocative efficiency - planners and economists have normally supposed that efficiency and

equity must be treated separately or that one or the other must dominate. Furthermore, analysts have normally preferred to focus on efficiency because it is much less controversial and less subject to variance among observers than is equity.

Finally, the case is made that, because poorer people or lower income areas usually are less efficient in producing goods than are their richer counterparts, the criteria for efficiency and quity are in basic conflict (cf. Mera, 1967). Where this is the case, alternatives serving the most needy areas or groups of people will not be among those alternatives which show highest net benefit in terms of contribution to economic growth. As a consequence of this normal conflict between quity and efficiency, tradeoffs between the two must be made in structuring the analysis and in the choice among alternatives by decision-makers.

Musgrave's landmark study (1959) argued effectively that public economics must treat both efficiency and equity, and proposed that these be considered as two branches of the subject. In practice, however, these considerations can not be separated, and somehow must be integrated in analysis and decision making. Many national and local programs explicitly have both efficiency and equity objectives.

Three basic approaches to dealing with efficiency and equity objectives are found in the prescriptive literature and in practice:

1. Favor one objective to the exclusion of the second;
2. Set the level of performance with respect to one objective as a constraint, then maximize for the second objective:
3. Formulate an explicit inter-personal utility function between efficiency and equity, which addresses directly the tradeoffs between the two.

Focus On Single Objective

In the first of these approaches, either efficiency is ignored and that alternative project or set of projects within the budget is chosen which best meets the equity criterion, or the most efficient choice is made regardless of the equity implications. This approach, which rejects the notion that decisions can serve both objectives, is the common case with benefit-cost analysis.

For example, a major work on public project selection by Eckstein (1958) reasons that projects should be ranked and selected on the basis of the size of their benefit-cost ratios until available resources for this type of project are exhausted. This is consistent with the position taken by many that the focus of economics should be allocative efficiency rather than a broader welfare economics perspective. As Baumol (1965: 356) argues:, "There is nothing in economic analysis which permits us to say that individual A should optimally receive (more net benefits than) B. The value judgments involved in recommending a distribution of income must somehow be grafted into the economic information ...".

An approach for 'grafting on' equity considerations is suggested by Lee (1983). This strategy places primary emphasis on the allocational aspects of public sector decisions, but recognizes the political and often legal necessity for dealing with distributional implications as well. Alternatives are seen as consisting of three categories of elements: characteristics of the investment, pricing for use of the investment as well as policies concerning its operation, and the means of financing that portion of the costs not met by user charges. It is Lee's argument that investment and the design of user charges, the first two categories, should concentrate

on maximizing net benefits, that is on seeking efficiency. The balance of the needed financing can then be designed to correct or adjust for undesirable equity impacts.

	efficiency	equity
outlay	Investment (cost)	
revenue	Pricing (user charges)	Residual Financing

Figure 1. **A Classification of Efficiency/Equity Elements**

User benefits and secondary benefits of many public investments, such as the major urban transportation projects or programs which are Lee's primary concern, predominately go to higher income populations. Lee's analysis is most appropriate for correcting the negative redistributional effects of these kinds of projects, rather than for seeking actions intended to result in 'progressive' transfers among social groups. Still, his framework for assessment highlights a number of major issues which require consideration in decision making.

When efficiency is the dominant criterion, user charges should be designed to eliminate consumer surplus, and inequity occurs when those benefiting do not bear the full costs. This form of inequity is difficult to correct through non-user fees to meet residual financing needs, because the instruments for accomplishing such transfers are crude and because of constraints on the political process in selecting and adjusting such instruments. Consequently, in this case, it is most desirable "... to match user charges and benefits as closely as possible" (Lee, 1983: 49).

There are two situations, according to Lee, that justify deviation from this procedure. First, services may be targeted to very specific groups, such as low-income or handicapped people, in which case residual financing would come from general revenue sources. Second, some services have impacts that are so local in nature that a special district tax can be used to capture the indirect benefits. Examples include local improvement districts paying for facilities that attract customers, and concessions from property owners who benefit from subway access. In this second situation where costs are not fully supported by users, indirect tax instruments that apply to the affected area meet both equity and efficiency goals better than do subsidies from general revenue sources.

One Objective As A Constraint

Use of the second approach involves setting either the efficiency or equity criterion at a minimum acceptable level, then selecting among alternatives on the basis of how well they perform with respect to the other criterion (Tabb, 1972). Thus, for example, candidate projects may be required to benefit primarily lower income families, and those alternatives meeting this constraint are compared on the basis of their relative efficiency. Treatment of one criterion as a constraint stops short of requiring that decision-makers specify their trade-offs between equity and efficiency.

Marglin (1967) has proposed that this approach be applied in an iterative manner, by inspecting the results of this procedure and then reevaluating the constraints. This would facilitate balancing the equity and efficiency concerns. Iterative ad-

justment of constraints in this manner appears to be most applicable when evaluating a small number of alternatives that are long range in nature.

Preference Function: Efficiency versus Equity

The third approach involves developing an objective function of some form which consolidates the social distribution and efficiency criteria for the purpose of selecting among planning alternatives. This implies something like a welfare function, development of which is a daunting undertaking, but additionally requires decision-makers to articulate how much efficiency they are willing to forego in meeting higher levels of their equity criteria.

A useful illustration of this approach is provided by McGuire and Garn (1972), in which they evaluate projects aimed at increasing employment in economically depressed areas for both their anticipated efficiency and equity outcomes. To do so, they construct an 'index of need' for various communities based on the exployment rate and median family income of these areas relative to national figures, and on the decision-maker's judgment of the relative importance of these two indicators of need.

This 'index of need' is interpreted as the marginal utility of benefits received by a community, which varies by community based on the welfare criteria used. Thus, in the case of a community with welfare criteria equal to the national averages, the index of need would be one (unity), and projects would compete for resources on the basis of efficiency alone. Each community's index of need is used as a weight to multiply the benefit-cost ratios for projects that would affect that community and, overall, that set of projects is selected which maximizes the sum of these project scores while exhausting available resources.

In reporting their application of this procedure, McGuire and Garn show that projects to communities most in need are not among the most efficient, and that while projects selected by their method sacrifice some efficiency, they are more efficient than a choice constrained by serving the most needy areas first. They conclude that an evaluation approach which treats explicitly the trade-offs between efficiency and equity, in this case through weighting benefit-cast ratios by a welfare index, best meets both criteria and presents the decision-maker with information concerning how much sacrifice in efficiency is involved in moving to choices that are more effective in meeting equity or need criteria.

In similar work that is more abstract but intended for application, Neidercorn (n.d.) develops a social welfare function based on consumption and leisure, and demonstrates the existence of a saddle point which maximizes production while minimizing inequality in the distribution of these resources. Equity is measured using a ratio of utility to deservingness, where the later is based on skill or marginal product per unit of time. If all workers were equally skilled, this model prescribes an equal distribition of resources to all. While this is but one definition of equity, the model may be generalized to include other measures of need.

Neidercorn's methodology is based on the work of Bergson (1938) and Lerner (1964), and shows promise for future development into a form that will be useful for decision making. In the meantime, methods such as the one proposed by McGuire and Garn are improvements over making judgments in structuring evaluation: e.g. that equity and efficiency are equally important, thus applying equal weights to the sets of criteria used to assess each of these objectives (Miller, 1980).

Most research addressing efficiency and equity goals falls into one of these three groupings. The difficulty of combining the two concerns analytically is apparent: The first two methods treat equity and efficiency separately, and the third involves developing a preference function that many analysts would rather avoid. Furthermore, we have extensive experience with assessing efficiency effects. We do not, however, have a comparable technology for dealing with the social distribution or equity effects of plans and policy. In the next sections, the reasons for this lagging technology are discussed, and several methods for measuring the distribitional implications of planning alternatives are explored.

WHY IS ANALYSIS OF EQUITY DIFFICULT ?

There are several reasons why developing rigorous techniques for assessing social equity implications has proven to be a complicated task. The first of these is the difficulty in specifying and measuring the kinds of societal resources which should be included in the evaluation. These resources include wealth and earnings and the material goods that these will purchase. Important non-material resources include esteem, participation and political influence, and social mobility (Miller **et al.**, 1970; 8), which are less tractable to analysis. For example Lineberry (1977) found an equal distribution of public services among areas and groups in the city that he investigated, but failed to assess the mobility and ability to exercise choice of various social groups.

A second difficulty is defining what constitutes an equitable or fair distribution of these resources. A range of definitions may be found in Blanchard (1983), Hochschild (1981), and Levy **et al.** (1974). These varying definitions are based on a variety of social beliefs and values, all held by one or another portion of the population, which thus makes it difficult to get agreement on a single definition. For example, Turner and Starnes observe that freedom from poverty is increasingly seen as a right of membership in a society (1976; 140-149), while Gans finds that work-ethic values are a major factor for people opposing transfers not based on personal effort (1973; 151-159). In short, any prescriptive definition of what constitutes a just distribution of resources is the subject of judgment.

A third difficulty is that since defining what constitutes social equity is a matter of judgment, it is appropriately a task of political decision making. This is illustrated by Lasswell's (1949) often cited definition of politics -"who gets, who pays"- which also points up the importance of distributional concerns in the political arena. For planners to address equity effects in their analysis is to invite controversy. It is heartening that planners are coming to accept the notion that conflicts on these grounds need to be revealed in order for negatiations to take place (Krumholz **et al.**, 1975; Krumholz, 1982), and that such conflict will serve to avoid evasive complacency (Miller and Roby, 1970).

Finally, a number of strategic questions pose difficulty. One of these questions concerns the effectiveness of effort spent in evaluating equity effects. Some argue that the costs of securing distributional information outweigh the benefits in terms of more equitable decisions. Others, especially those employing a Marxist perspective, argue that such analysis is counter-revolutionary because it diverts effort from working to alter the present economic order (Harvey, 1973: 144-145). Still others seek to demonstrate that focusing attention on the distributional effects of proposals does influence public decisions
(Levy et al., 1974).

Another strategic question is whether the focus should be on vertical equity or on horizontal equity. Vertical equity is concerned with transfers among income classes or to disadvantaged groups. Many planners subscibe to the position articulated by Rawls (1972): that 'social primary goods' should be equally distributed, with any inequality favoring the least advantaged. Horizontal equity refers to transfers between individuals and groups within the same income groups, among political jurisdications or geographic areas, users and non-users, etc. Undesirable redistribitions of this sort may result from public actions, and factual analysis both can diminish controversy based on misinformation and can provide the basis for negotiation.

A final strategic question which poses difficulty concerns how to deal with the time dimension. This commonly takes the form of how much to invest today for the benefit of future generations (Baumol, 1968; d'Arge **et al.**, 1982). It also takes the causal form of whether aid to low income persons serves to institutionalize poverty and to diminish personal efforts at self-help.
These several difficulties in dealing with distributional issues help to explain why many planners and policy analysts have avoided or neglected to evaluate the 'social equity' implications of public actions. Consequently, as Wildavsky (1979) points out, distributional outcomes are seldom decided upon but rather result from decisions and other forces. Yet, most public actions do have redistribitional effects, these effects are commonly enormous (Tabb, 1972), improved social equity is a purpose of many public programs, these effects are of concern to political decision-makers, and planners are expected to treat equity in an explicit manner as a part of their professional responsibility (A.I.C.P., 1980; Clavel and Goldsmith, 1970). How then can we address the distributional implications of planning alternatives in evaluation ?

MULTICRITERIA METHODS FOR MEASURING SOCIAL EQUITY EFFECTS

Even though the social equity effects of public actions are regarded as important by planners and others involved in the decision-making process, there are few examples of these effects being measured in practice. Since decision making is most responsive to those factors that we measure, it follows that the distributional implications of plans and policies are not going to be treated adequately in decision making until these implications are measured as a routine part of evaluation.

The problem of measuring equity effects is addressed by three promising methodological approaches. The first of these assesses the degree to which affected groups of people share equally in what they regard to be the benefits of planning alternatives. The second approach involves the familiar weighted-sum method of evaluation, but retains the weights of various groups in a disaggregated form, and thus provides a set of scores representing the views of each group. The third approach avoids summarizing of performance scores, presenting instead simple incidence profiles for each affected group.

Each of these methods explicitely display evidence concerning the distributional implications of alternatives, but they vary in their information requirements and the extent to which they lead to recommendations. These three methods will be discussed in greater detail and applications will be used to illustrate each.

A Multiple Attribute Method for Assessing Equality of Benefits

This method addresses how various designs for services or land uses differentially

benefit affected groups of people (Miller, 1981). Each planning alternative is analyzed as having a set of attributes or dimensions that contribute to the value received from that alternative (Lancaster, 1966; Lancaster, 1971; Miller, 1974). Each group of people included in the evaluation views the relative importance of these attributes differently, in terms of how these features of the good or service contribute to their satisfaction. A weighted sum approach (Hill, 1968, 1973; Miller, 1980; Voogd, 1983), which combines the group specific weights with the related attribute scores, is used to indicate the benefit or satisfaction received by each group. These summary scores are compared to assess the degree to which each affected group is equally benefited.

By way of example, attributes for urban playfields include the number of playing grounds, their condition such as drainage and surface, availability of night lighting, whether they could be reserved and their availability, hours at which they are open, etc. Group specific weights, indicating the value of each of these attributes, may be treated as constants if attribute scores vary over a narrow range. The average satisfaction that each group receives from available playfields may be estimated by the equation:

$$S_g^r = \left(\sum_{j=1}^{n} \sum_{k=1}^{n} w_g^k a_j^k \right) / n ,$$

where:

S_g^r = the average satisfaction for group g from n sources of service r;

w_g^k = the priority weights of group g for attribute k, where k=1...m;

a_j^k = the estimated v alue of attribute k for this service from source j (j=1...n).

This procedure is a variation on the goals achievement form of multicriteria evaluation, with sources of a service as alternatives represented as columns, and attributes of the service substituted for criteria and represented as rows. Performance scores for each attribute are standardized to provide a common metric, and the sum of the products of these and the valuation weights for a group provide a measure of the satisfaction received by that group from that source.

When several types of services are being evaluated, and several groups are involved, the following table provides a useful way of presenting the results:

groups	services r_1	r_2	r_3	(% of pop.)
g_1				
g_2		S_i^j		
g_3				

Figure 2. **The presentation of distributional data**

This illustration includes three services r_j (j=1,2,3) and three population groups g_i

($i=1,2,3$). The S_i^j are weighted average satisfaction scores for each group for each type of service. Comparision of these scores is aided by referring to the relative contribution by each group to the population as a whole, which appears in the row stub of the table.
This method also can be applied to assess the distribitional implications of the way that a set of services is being provided currently. We are in the process of applying this method **ex ante** to estimate the distributional implications of alternative project proposals for additional organized-sport playfield capacity in Seattle. In this case, because access is an important attribute, the city has been subdivided into analysis zones, and tables of the form shown in Figure 1 are prepared for each of these zones, with each package of alternative projects being represented as a column in these tables (Turner, 1972; Miller, 1981). Since attributes such as distance, user fees, and waiting time are 'cost criteria' (Voogd, 1983: 79), normalized scores for these are subtracted from one, so that larger scores denote greater satisfaction.

User groups are identified from records and population information, and subdivided into groups on the basis of socio-economic variables. These groups are represented as rows. By means of a survey, representatives of each group are asked to supply weighting information for the service attributes, which are referred to as 'service features'. Each informant is provided with a list of fourteen service features, stated in terms of average performance scores for the currently available set of facilities. They are asked first to rank service features that they would like to see improved, then to distribute 100 points over these service features. Preliminary results suggest that respondents had little difficulty with this task.

The packages of alternative proposals in each case include existing facilities, and are differentiated by alternative sets of improvements to existing facilities, and provision of new facilities: some at newly acquired sites, some as joint use of school grounds. The weighted sum for each group for each alternative will be compared with the counterpart numbers for other groups to provide information on the distributional impacts of these alternatives.

A similar weighting procedure has become a routine part of neighborhood planning in Seattle over the last several years. In this application, a representative advisory body participates in developing a physical development plan for the neighborhood, including proposed land use changes and zoning, traffic and street improvements, and parks and other facilities. One result is a list of desired capital improvements, each with an associated price.

A survey in a ballot format, listing possible capital improvements, is sent to all residents of the neighborhood. This survey instrument includes a set of stickers with dollar values printed on them. Each respondent is asked to allocate these stickers over the list of alternative capital improvements, which requires making choices since the budget represented by the supply of stickers is smaller than the cost of the full list of possible improvements. Self reported socio-economic information for the household permits using these ballots for assessing the distributional implications of various sets of capital improvements.

Variations of this method are useful in many planning exercises. For example, Breheny (1974) assessed the spatial opportunity implications of several land use configurations by calculating the mean distances to various numbers of jobs and to other people, to various levels of shopping, and to other activities. When combined with information concerning the spatial distribution of various population groups, the equity implications of land use planning alternatives may be assessed.

A major feature of this method is its analysis of the multidimensional nature of a public service or set of services, how each attribute or dimension of a service contributes to the value received by clients of that service, and how various client groups differ in their valuation of these dimensions. Thus this method estimates the real income effects of alternative ways of providing a service, alternative configurations of facilities, or even alternative components of a physical development plan.

Equally effective provision of a service to all groups would result in equal summary figures for each group. Variation among these figures, or relative deviation from some base point figure such as the smallest or the average score, demonstrates the extent to which these groups do not receive equal benefits from an existing situation or from a planning alternative under consideration.

With spatially disaggregated analysis, in which the planning area is subdivided into analysis zones, bar graphs of the results may be presented as a third dimension to a map base, showing how well members of a group are served depending on where they live (Blanchard and Miller, 1981). The number of data items required by this method are modest, since only one set of attribute scores for each source of a service is needed except in the case of accessibility measures for a set of analysis zones, and it is usually reasonable to use one set of priority weights for each client group for all sources of a service.

Group Specific Assessment of Planning Alternatives

Conventional multicriteria evaluation employs a matrix of planning alternatives represented as columns and criteria represented as rows. All entries, which are performance scores of each alternative with respect to each criterion, are multiplied by a weight which accounts for the relative importance of that criterion, and these products are summed to provide an overall effectiveness score. These weights are usually an average for the affected population.

This second method for dealing with the distributional effects of planning alternatives is simular to the conventional approach just described. However, each group included in the evaluation is represented by its own set of weights for the criteria employed. Thus the set of weights for each of several groups form a matrix. The dot product of this weights-by-criteria matrix, and the effectiveness or criteria-by-alternatives matrix, produces a rectangular matrix of summary scores in which each of the groups is represented by a row of normalized, weighted effectiveness scores for the set of planning alternatives. This is similar to the 'appraisal matrix' discussed by Voogd (1983: 64), and takes the form:

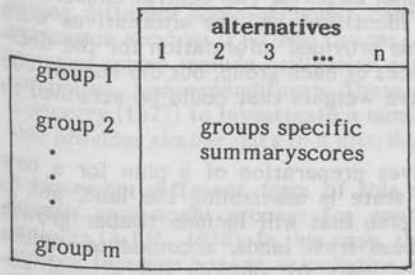


Figure 3. **Summary Scores**

Inspection of this summary score matrix, by rows, reveals that alternative favored by each group based on the information included in the evaluation, but does not provide a single summary score for the overall effectiveness of each alternative. While this method accounts for the differing views of various affected parties, it provides only partial and often indirect evidence concerning the impacts of each of the alternatives on each of the groups, since the raw performance scores are not disaggregated by incidence group. In many decision-making contexts, analysis of the preferences of several groups is the kind of information that is sought. Elected officials commonly undertake this sort of assessment, though in an unanalytical and general manner, to understand the reactions of their constituents.

Groups included in this form of evalution may be defined on the basis of various shared interests or self identities. Where vertical equity is the issue, the groups are appropriately defined on the basis of personal or family income, and it may be desirable to summarize the group-specific scores for each alternative into a single score. In this case, a social equity weight will need to be developed for each group, to represent the relative social importance given the preferences (or benefit) of that group.

Such an equity weight must account for two variables: a fairness criterion which is applied to each person based on group membership, and the proportional contribution of each group to the total population. While the second of these is easily calculated, the fairness criteria is based on a social norm or definition of equity. As we have seen, there is no consensus concerning what is a socially just distribution of resources. However, the marginal income tax rate for each income group, while seldom the result of a deliberate policy-making process, does constitute a de facto societal definition of equity. This marginal tax rate, when subtracted from one, provides a measure for this definition of equity which is usually more progressive in its redistributional implications than are the effects of many public investment decisions (Piven and Cloward, 1977).

The social equity weight for each group is the product of the fairness criterion and the percent of the population in that income group. Each group's summary score for each alternative may be multiplied by the group's equity weight, these products may be displayed in a new matrix of equity-weighted scores, and these scores may be examined to assess the distributional implications of the alternatives. If desired, these equity-weighted scores for each alternative can be summed to provide an overall score for comparison with similar scores for the other alternatives.
A Dutch study of alternative airport locations (Ministerie van Verkeer en Waterstaat, 1979), provides a partial illustration of this method. Alternatives included possible locations for a second national airport and schemes for decentralizing this function to several regional airports. The central concerns of various interested or affected parties were identified, and the alternatives were assessed from each of these perspectives. This provided information for the decision makers concerning the relative preferences of each group, but did not include the more normative step of developing relative weights that could be attached to the views of each group.

A simular study that is just underway involves preparation of a plan for a new state forest in the vicinity of Seattle. The state is assembling the land, and is developing a multiple purpose management plan that will include timber growth and harvesting as a revenue source from these trust lands, accomodate various recreational activities, and provide an opportunity for outdoor educational activities. An advisory panel has been established, on which are represented local governments, the timber industry, and a variety of educational and recreational

groups who have a stake in the plan. The activities of several of these groups can conflict; for example trail-bikers, hikers, hang-gliders, and timber harvesting.

We expect to assess the relative attractiveness of alternative plans from the viewpoints of each of these groups. The advisory panel will assist in developing this plan over a period of about a year and a half. Consequently, members will be very familiar with the issues involved, and in continuing contact with their respective organizations. It is reasonable to expect that these panel members will be informed and willing sources of information concerning the relative importance of criteria from the viewpoint of their respective groups.

The major result of this analysis will be a matrix of the evaluations of alternatives from these viewpoints. While the relative sizes of these interest groups will be estimated, it is unlikely that this information will be used to weight the groups in order to calculate an overall summary score for each alternative. The results of this analysis will be presented to the panel for its discussion, and to the state agency involved, as evidence to inform its final decision.

Unsummarized Displays of Distribution Effects

The third method of evaluating planning alternatives for their equity implications employs a simple accounting framework. A table is developed, in which alternatives are represented by columns and impacts by rows, but the benefits and disbenefits of each alternative are disaggregated by incidence group. Incidence profiles of this sort are similar in format to the familiar 'Planning Balance Sheet' developed by Lichfield (1966, 1969) and to Quade's (1975) 'Score Card'.

Another format, displaying the same information, employs a separate matrix or table for each of the affected groups (Thomas and Schafer, 1970: 62-63). Rows for each group in the case of the single-table format are consolidated into a separate table for each group. When several groups are included in the analysis, the separate table for each group tends to be more difficult to interpret and use than does the single table of distributed impacts.

This simple information display method is recommended by the guidelines to applicants for federal grant assistance issued by the U.S. Urban Mass Transit Administration (1977), in meeting the U.S. Civil Right Act of 1964. The information that is required describes the distribution of transit services by ethnic groups, and is primarily drawn from census data.

Another illustration of this method evaluates whether two areas with populations having different socio-economic characteristics share equally in the provision of recreation services (Fisk and Lancer, 1974). Measures employed and displayed in a series of tables include utilization rates, quality and quantity of recreational opportunites, and expenditures. These measures are similar to the ones used by Lineberry (1977) to investigate a number of urban public services. Weisbrod (1968) also provides similar data in a distributional study of investments for recreation.

A somewhat different form of this method is employed in the Simi Valley net benefit assessment process for evaluating single public or private investment projects (A.I.P., 1978). In this case, the project is represented by a table in which project features appear as columns and affected groups as rows. All entries are ordinal level estimates of the benefit that each group receives from each project feature. Notes, keyed to the larger of these effects, describe the basis for these estimates and sometimes their magnitude. Mitigating measures for disbenefits are

described in the row stub of the table and represent marginally different alternatives to the central proposal.

A final example of incidence profiles in the form of a score card or balance sheet is provided by Schaeman and Muller (1974:33-34). Various kinds of impacts are enumerated, measures for each are identified and, for a particular project proposal, ratio level estimates of these impacts are made for specified 'clientele groups' including the populations of the immediate neighborhood, of the jurisdiction, and of the low-income families in the jurisdiction.

Each of these examples present estimates of the distribitional effects of planning alternatives or of a single proposal in the form of an account, in a manner similar to the familiar environmental impact statement or report. Decision-makers are left to their own devices to assess and interpret this listing of information. Various types of impacts usually are estimated using a variety of measures. Consequently, the natural tendency to summarize across impacts on each group is precluded. These performance scores could be normalized, as was done by the first two methods that were discussed. Even so, none of the examples of performance profiles that were reviewed include information concerning the relative importance to the affected groups of the various impacts. At least in the U.S., decision-makers commonly demand recommendations and 'bottom line' figures from analysts, and are impatient with information displays which do not include these. Summary scores could be calculated for these examples by employing group specific weights for the set of impacts assessed, as in the first two methods. These summary scores would provide estimates of the marginal social value of redistribution by incidence group.

SUMMARY

Increasing social, political, and legal demand that both efficiency and equity effects be considered has added to the complexity of decision making. Two major approaches to dealing with both of these objectives are available. One approach treats one of these objectives as a constraint and seeks to maximize the other. Sometimes this is done iteratively in an effort to reassess the constraint. This approach appears to be most useful for evaluating a small number of long-term investments.

The second approach seeks to deal with both objectives simultaneously, and to confront the problems of trade-offs between them. This line of development is dealing with relatively new ground, since it involves constructing and applying a welfare function and must deal with the inter-personal comparison of utilities. While these are concepts that are seen widely in economic theory, they are largely untried in practice.

There is still considerable uncertainty about the appropriateness of methods used in each of these two approaches. Yet, especially with increasing recognition that resources are scarce, there is demand for reliable evaluation that adequately treats both the distributional and efficiency effects of planning and policy alternatives.

As we have seen, the problem does not end with finding acceptable means for analytically balancing between efficiency and equity. The definition of what constitutes an equitable distribution of resources is a matter of judgment on which there is no consensus. Additional problems in dealing with equity further explain why analysts who wish to avoid controversy have omitted it from their evaluation

work, emphasizing efficiency instead.

As a consequence, there is still much to do in developing methods for validly assessing the distributional implications of planning alternatives. Three promising methods for doing this are discussed, each presenting different information for decision making. More work on methods of these kinds is needed. The increasing complexity of decision making, and the resulting greater dependence on technical analysis and advice, is generating demand that should facilitate further methodological development for evaluating the distribution, and efficiency, effects of investment decisions.

REFERENCES

American Institute of Certified Planners, Code of Professional Responsibility and Rules of Procedure, 1980 Roster, A.I.C.P., Washington (1980)

American Institute of Planners, The City of Simi Valley, California - Net Benefit Assessment Process, American Institute of Planners, Washington (1978)

Baumol, W.J., **Economic Theory and Operations Analysis**, Prentice-Hall, Englewood Cliffs, New Jersey (1965) 2nd ed.

Baumol, W.J., On the Social Rate of Discount, **American Economic Review**, 58, 4 (1968) 788-802

Bergson, A., A Reformulation of Certain Aspects of Welfare Economics, **Quarterly Journal of Economics**, 52 (1938) 310-334

Blanchard, W.H. and D. Miller, Methods of Measuring the Distributional Effects of Plans and Projects, Paper presented to the Annual Meeting of the Western Regional Science Association, Monterey, California (1980)

Blanchard, W.H., Social Equity and Planning Analysis: How Planners Can Think and Talk About Fairness, Unpublished Ph.D. thesis, Department of Urban Planning, University of Washington, Seattle (1983)

Braybrooke, D. and C.E. Lindblom, The Problem of Distribution, In: D. Braybrooke and C.E. Lindblom, **A Strategy of Decision - Policy Evaluation as a Social Process**, Free Press, New York (1963) 169-199

Breheny, M.J., Towards Measures of Spatial Opportunity, **Progress in Planning**, 2, 2 (1974) 97-105

Clavel, P. and W.W. Goldsmith, Distributions and Redistributive Processes: Two Foci For Planning Education and Practice, Paper presented to the October 1970 Conference of the American Institute of Planners, Minneapolis - St. Paul (1970)

D'Arge, R.C., W.D. Schulze and D.S. Brookshire, Carbon Dioxide and Intergenerational Choice, **AEA Papers and Proceedings**, 72, 2 (1982) 251-256

Eckstein, O., **Water Resource Development: The Economics of Project Evaluation**, Harvard University Press, Cambridge MA. (1958)

Fisk, D.H. and C.A. Lancer, **Equality of Distribution of Recreation Services: A**

Case Study of Washington D.C., The Urban Institute, Washington (1974)

Gans, H., **More Equality**, Vintage Books, New York (1973)

Harvey, D., **Social Justice and the City**, Johns Hopkins University Press, Baltimore Md. (1973)

Hill, M., A Goals-Achievement Matrix for Evaluating Alternative Plans, **Journal of the American Institute of Planners**, 34, 1 (1968) 19-28

Hill, M., **Planning for Multiple Objectives**, Regional Science Research Institute, Monograph Series No. 5, Philadelphia (1973)

Hoschschild, J.L., **What's Fair: American Beliefs About Distributive Justice**, Harvard University Press, Cambridge Mass. (1981)
Horowitz, I.R., Methods and Strategies in Evaluating Equity Research, **Social Indicators Research**, 6 (1979) 1-22

Krumholz, N., J.M. Cogger and J.H. Linner, The Cleveland Policy Planning Report, **Journal of the American Planning Association**, 41, 5 (1982) 298-304

Krumholz, N., A Retrospective View of Equity Planning - Cleveland 1969-1979, **Journal of the American Planning Association**, 48, 2 (1982) 163-174

Lasswell, H.D., **The Analysis of Political Behavior**, Oxford University Press, London (1949)

Lancaster, K.J., A New Approach to Consumer Theory, **Journal of Political Economy**, 74 (1966) 132-157

Lancaster, K.J., **Consumer Demand: A New Approach**, Columbia University Press (1971)

Lee, D.B., Technical Guidance for Project Planning Evaluation of Urban Transportation Capital Investment Projects, Staff Study, Research and Special Programs Administration, U.S. Department of Transportation, Cambridge, MA (1983)

Lerner, A.P., **The Economics of Control**, MacMillan, New York (1944)

Levy, F., A.J. Meltsner, and A. Wildavsky, **Urban Outcomes**, University of California Press (1974)

Lineberry, R.L., **Equality and Urban Policy: The Distribution of Municipal Public Services**, Sage Publications, Beverly Hills, CA. (1977)

Lichfield, N., Cost Benefit Analysis in Urban Redevelopment - A Case Study: Swanley, **Urban Studies**, 3 (1966) 215-249

Lichfield, N., Cost Benefit Analysis in Urban Expansion - A Case Study: Peterborough, **Regional Studies**, 3, 2 (1969) 123-155

Marglin, S.A., **Public Investment Criteria**, M.I.T. Press, Cambridge, MA. (1967)

McGuire, M.C. and H.A. Garn, The Integration of Equity and Efficiency Criteria in Public Project Selection. In: K.E. Boulding and M. Pfaff (eds.), **Redistribution to**

the Rich and the Poor, Wadsworth Publishing Company, Belmont, CA. (1972) 357-369

Mera, K., Tradeoff Between Aggregate Efficiency and Inter-regional Equity: A Static Analysis, **The Quarterly Journal of Economics** (1967) 658-674

Miller, D.H., **Community Mental Health: A Study of Services and Clients,** D.C. Heath, Lexington, Mass. (1974)

Miller, D.H., Project Location Analysis Using the Goals Achievement Method of Evaluation, **Journal of the American Planning Association,** 46, 2 (1980) 195-208

Miller, D.H., Plans and Publics: Assessing Distribution Effects, Planologisch Memorandum 81-2, Dept. of Civil Engineering, Delft University of Technology (1981)

Miller, S.M. M. Rain, P. Roby, and B.M. Gross, A Social Indicators Definition, In: R.E. Will and H.G. Vater (eds.), **Poverty in Affluence,** Harcourt, Brace and World, New York (1970), 18-22

Ministerie van Verkeer en Waterstaat, **Structuurschema Burgerluchtvaartterreinen,** Staatsuitgeverij, The Hague (1979)

Niedercorn, J.H., Equality, Pareto Optimal Utility and the General Linear Social Welfare Function, Report of the Modeling Research Group, Department of Economics, University of Southern California, Los Angeles (no date)

Piven, F. and R. Cloward, **Poor Peoples' Movements: Why They Succeed, How They Fail,** Vintage Books, New York (1977)

Quade, E.S., **Analysis for Public Decisions,** American Elsevier, New York (1975)

Rawls, J., **A Theory of Justice,** Harvard University Press, Cambridge, Mass. (1971)

Schaenman, P.S. and T. Muller, **Measuring Impacts of Land Development,** The Urban Institute, Washington (1974)

Tabb, W.K., Alternative Futures and Distributional Planning, **Journal of the American Institute of Planners,** 38, 1 (1972) 25-32

Thomas, E.N. and J.L. Schafer, Strategies for the Evaluation of Alternative Transportation Plans, **National Cooperative Highway Research Report,** 96, Highway Research Board, Washington (1970)

Turner, C.C., A Model Framework for Transportation and Community Plan Analysis, **Journal of the American Institute of Planners,** 38, 5 (1972) 325-331

Turner, J.H. and C.E. Starnes, **Inequality; Privilege and Poverty in America,** Goodyear Publishing, Santa Monica, CA. (1976)

Urban Mass Transportation Administration, Circular UMTA C1160.1, Subject: Interim Guidelines for Title VI Information Specific to UMTA programs, U.S. Department of Transportation, Washington (1977)

Voogd, H., **Multicriteria Evaluation for Urban and Regional Planning,** Pion, London (1983)

Webber, M.M., Planning in an Environment of Change, **The Town Planning Review**, 39, 4 (1969) 277-295

Weisbrod, B.A., Income Redistribution Effects and Benefit-Cost Analysis, In: S.B. Chase Jr. (ed.), **Problems in Public Expenditure Analysis**, The Brooking Institution, Washington (1968), 177-209

Wildavsky, A., The Political Economy of Efficiency: Cost Benefit Analysis, Systems Analysis, and Program Budgeting. **Public Administration Review**, 26 (1966) 292-310

Wildavsky, A., **Speaking Truth To Power: The Art and Craft of Policy Analysis**, Little, Brown and Company, Boston (1979)

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

FROM IMPACT ASSESSMENT TO IMPACT EVALUATION

N. Lichfield
Nathaniel Lichfield and Partners, London
Intern. Centre for Land Policy Studies, Cambridge, Mass.

1. THE GROWING COMPLEXITY OF THE PUBLIC DECISION-MAKING CONTEXT

In the reaction against the growth of central direction in societies, both east and west and north and south, there has been a re-emphasis of the role of the market. The attraction here is that many decisions which otherwise remain to be made by Goverment can emerge form the "unseen hand" of the interplay of individual smaller scale decision centres.

But while there are clear attractions in this counteremphasis, there is no claim, even by Hayek or Friedman, that a rejuvenated market system can replace entirely the role of decision taking by representative government. And indeed in one of the more striking attempts at reversal towards the market, contemporary Britian, the very philosophy of "less government" has currently produced prospective legislation aiming at controlling local government expenditure, which would introduce far greater central control of local government freedom to devise local progammes than has ever existed before in Britain. (1) If it materialises the new legislation would give to central government a role in public decision making which could be both decisive in the extent and nature af local government programmes, and also need to establish some balance between what occurs in localities themselves. The "democratic" isues are controversial; government supporters, in both local and central government, including former Ministers, are opposing the Government.

Thus despite any attempts at "liberalisation" of decision taking away from government, there is inevitably a very large core of central decision taking. This is being extended to international government, in the various combinations that exist and are emerging: in the United Nations, Common Market, GATT, etc. But not only are we faced with this persistent and enlarging core, but it is becoming more complex if only for one reason: the increasing complexity of the institutions involved in any particular kind of decision making, which I contrast with decision taking, for which it is the context. This is probably the result of many trends, such as the search of bureaucracy to find means of decision making which must reflect the complexities of modern government; the increasing awareness and involvement of the public in affairs, which require that the bureaucracies be sensitive to an increasing number of views and considerations from outside the departments themselves; and the increasing diversity and complexity of the institutions which need to be involved in the decision making process.
All this means that the focus of this workshop, on complex policy problems in the public sector, is not likely to be of dim°inishing importance in the foreseeable future. And indeed this could be the one constancy in the uncertain world, namely the imperative for public sector bodies to continue to make decisions on behalf of

their constituencies despite the difficulties of so doing. This being so, the contribution of evaluation methods and techniques in the reaching of such decisions cannot aim at the " optimal". It must have the more humble and realistic "satisficing" criterion of : how can we attempt to sensure that the decision taken with the aid of these methods and techniques is a better one than would otherwise have prevailed?

2. MEETING THE DECISION-TAKERS CRITERIA

If this objective for evaluation methods is to be met then clearly there can be "... no general consensus about which kind of evaluation method is most appropriate ..." (2) since the choice of evaluation method should be oriented to the decision making problem, organisation and constraints with which the evaluation analyst is confronted. As put on another occation in Holland, following a review of the various evaluation methods, "It is not appropriate to ask which method is most suitable since, as noted, they are in the main designed to answer different questions". (3)

In order to assess this prior consideration we are faced with a series of questions which must provide the context for the evaluation analysis and recommendations. A possible list is: (4)

1. Does the evaluation relate to policy, plan, project or programme of projects ?

2. In respect of any of the preceding, what kind of planning process is envisaged ?

3. Where and how would the evaluation fit into this planning process ?

4. What role would others than the particilar decision takers have in the process (e.g. higher or lower level governmental bodies in the hierarchy, agencies to be consulted, the public at large).

5. What are the options before the decision makers, both as perceived by them and as might be added ?

6. What constraints are they imposing on the choice between options, which would rule out particular options ?

7. What are the decision taking criteria which they will be using, as for example minimising the cost to the public authority; securing maximum votes in a forthcoming election; maximising the public interest irrespective of the constituency limitations ?

8. What evaluation methods and techniques do the above indicate ?

9. What data are readily available without further research and investigation ?

10. When must the decision in itself be made, thus leading to the time constraints on the evaluation contribution ?

11. What uses are to be made of the evaluation analysis other than for the decisions, e.g. for public participation ?

12. What are the budget limitations on the evaluation analysis ?

All the preceding would lead to a conclusion on how best to frame an evaluation analysis which will best meet the requirements, namely aiding the decision makers in the best possible way within the limitations of time, resources, etc. One example from a current study will illustrate. (5)

The terms of reference asked for a test of the hypothesis that the conventional methods of evaluation for public transport did not reflect in full the costs and benefits of public transport to the community. Accordingly, following a discussion of principle which tested the hypothesis, there was a demonstration by application to a case study of the large industrial estate in Manchester, Trafford Park.

The methods were:

(1) **Conventional**

(a) corporate objectives achievement (COA), with the objectives being those hel;d corporately by the decision taker, namely Greater Manchester Passenger Transport Executive;

(b) financial costs and returns (FCR), being the method imposed in the statute governing the operations of the Transport Executive;

(c) social cost benefit analysis (SCBA), being the traditional method of economic evaluation officially required under the Public Transport Act of by the relevant government department, Department of Transport.

(2) **Non-conventional**

(d) framework appraisal (FA), being the method of appraisal for trunk roads by the same Department, which has been introduced since 1979 to extend the evaluation method of highway route selection beyond conventional cost benefit analysis;

(e) community impact evaluation (CIE), being the attempt to relate the operations of the Passenger Transport Executive to the development planning objectives of the Greater Manchester Council, which is the development plan making authority and which, as Passenger Transport Authority, gives guidelines within the Plan's wider objectives to the Transport Executive.

The three conventional methods related to the single sector of transportation and were fully numerical. They each showed a different ranking from amongst eleven options. The two non-conventional methods were multi-sector and only, for that reason, partly numerical. They showed a different ranking for the three options chosen from the eleven.

In essence therefore it is the choice of **method** which decides the choice of option to be implemented. And therefore the choice of method must reflect the constraints and criteria which the decision makers wish to raise, for otherwise they are not getting the choice they would really favour.

3. NEED FOR GREATER SYNTHESIS IN DECISION TAKING METHODS AND TECHNIQUES

The burden of the preceding section is the need to select from the available methods and techniques that which would be specifically tailored to the decision takers requirements in their particular context. Or it can be seen as the constraints obtained from the actual situation on the particular method of evaluation which has been pre-selected by the analyst, according to his professional background, preferences amongst the methods and experience in their applications. But even if the methods and techniques are usefully tailored in this way, there still is a significant step that could ne taken by the analyst pursuing the objective of helping the decision makers. This is to recognise that evaluation analysis is often hampered by methods being treated as discrete and isolated from each other, whereas in the array of methods which are currently available there are features from certain methods which can well contribute in the application of others. Some examples will illustrate.

(1) In the study of public transport just referred to the five methods of evaluations, on the surface, are widely disparate. But in fact they have many common features. Thus instead of pursuing the methods quite discretely they were seen as "nesting" within the widest ranging of the methods in terms of sectors considered, i.e. CIE; framework appraisal, while earning the title of "multisector" as against the others which were "single-sector", nonetheless did not range as widely as the community impact evaluation.

(2) Another instance of discreteness and failure to co-ordinate emerged at the earlier conference in Holland cited above (3). Summarising the discussion, Faludi referred to the papers presented at the Colloquim on strategic choice by Hickling and planning balance sheet analysis by Lichfield in the following terms: "... it has been a source of wonder why the strategic choice approach and Lichfield's approach to evaluation have never been more fully integrated in the past". (6) In the event this reconciliation has not taken place in the intervening years (emphasising the point of separateness mentioned above) but could be advanced at this present workshop.

(3) There is also the need for synthesis in these two methods of a different character than the one just cited. Planning balance sheet analysis was initially formulated as a method of project and plan evaluation, which was practised as a discrete step in the planning process. But the limitations of this approach were seen and in consequence the needs and methods of integration of evaluation into the planning process were studied. (7) In essence this shows how the evaluation approach, method and technique can and should be sustained throughout the plan making and implementation process.

By contrast, strategic choice, despite its title which suggests selection following evaluation, "... provides a general framework for public planning and also includes several techniques adapted to that framework ... but little attention has been paid so far to evaluation as a distinct step, or phase, in this strategic choice process. The workshop will attempt to explore the role of evaluation in strategic choice, therefore. It would transpire that this lack of attention is no accident but reflects a distinct view of the place of evaluation". (8)

(4) The final example comes from the comparison between the goals/objectives matrix approach and that of planning balance sheet analysis. Following the attempt to compare and reconcile the differences between these two approaches the following conclusion emerged:

> "Thus the GAM seems to be directed at examining whether the plans have achieved certain aims which the planners and decision takers consciously set out to achieve. The PBS is directed at something different: what will be the consequences of the plans in question for the welfare of those who are affected". (9)

Bases on this conclusion, the evaluation approach adopted by the writer has been to recognise that a prior step in the presentation of the plans, policies, etc. for evaluation is the **testing** of those plans in a variety of ways, short of evaluation proper as understood in a planning balance sheet analyses/ community impact evaluation, which is derived from economics as the comparison of options in terms of the outputs (benefits) as against the input (costs) to the whole community. (10) Thus the goals/achievements matrix approach is seen by the writer as a **test** of the question: to what degree have the objectives for the planning study been met in the planning solution. Accordingly the approach is to carry out this test as appropriate prior to the evaluation proper, as with other tests, such as political or economic feasibility.

Thus the general need that seems to be emerging is not so much the search for a consensus on "which kind of evaluation method is most appropriate" as a comparative review of the methods to see where each is **best** used (2 above) and the degree to which they can be used to sustain each other (3).

This approach has been explored by the writer in relation to the field of impact assessment which has grown up alongside the more generally understood methods of evaluation. How can the two support each other ? This is the topic in the remainder of this paper, with the indication of how the writer has sought to absorb impact assessment into community impact evaluation.

4. FROM IMPACT ASSESSMENT TO IMPACT EVALUATION

4.1. Origins of Impact Assessment

A review of the origins of impact assessment is particularly taxing,because of its diverse origins, the volume of work that hast been carried out since around 1970 and the rapid spread of practice around the world. (11)

The pioneer was probably in economic impact analysis, carried out as part of project appraisal, where the concern was with the changes in outputs of goods, services, income and employment from investment, be this in new buildings or plants, water resources, etc. In parallel, but of later origin and less developed, came impact analysis relating to social concerns, such as the growth of tourism in the developed or developing world, or the effects of urban revewal on established communities through break up of families and groups, displacement from job opportunities, etc. Under a different banner came the transportation impact which sought to predict the effects on urban and regional development activities of the introduction of new major tranportation facilities (as in the closing of railway lines). A fourth strand is the environmental impact whose origins were in applying the knowledge of ecologists, biologists, chemists and others to ascertain the environmental pollution consequences on natural resources, of the earth, plant and animal life. And finally are the more recent endeavours under the name of urban impact analysis which seek to predict the impacts of fiscal and economic measures which are not urban in themselves on the cities as urban entities, and on geographical and social sub-divisions in the cities. (12)

This growth of impact analysis and assessment has been explored along different paths by different skills and concentrated in particular professional offices or univesity departments. Inevitably the different streams touch upon each other. It is difficult to be concerned with social impacts without recognising that income and employment are important elements in social well-being, just as it is difficult to be concerned with economic impact without recognising that welfare is not totally subsumed in income and employment. Thus these apparently independent streams have the propensity to mutually stimulate each other, just as the growth of transportation modelling stimulated land use modelling.

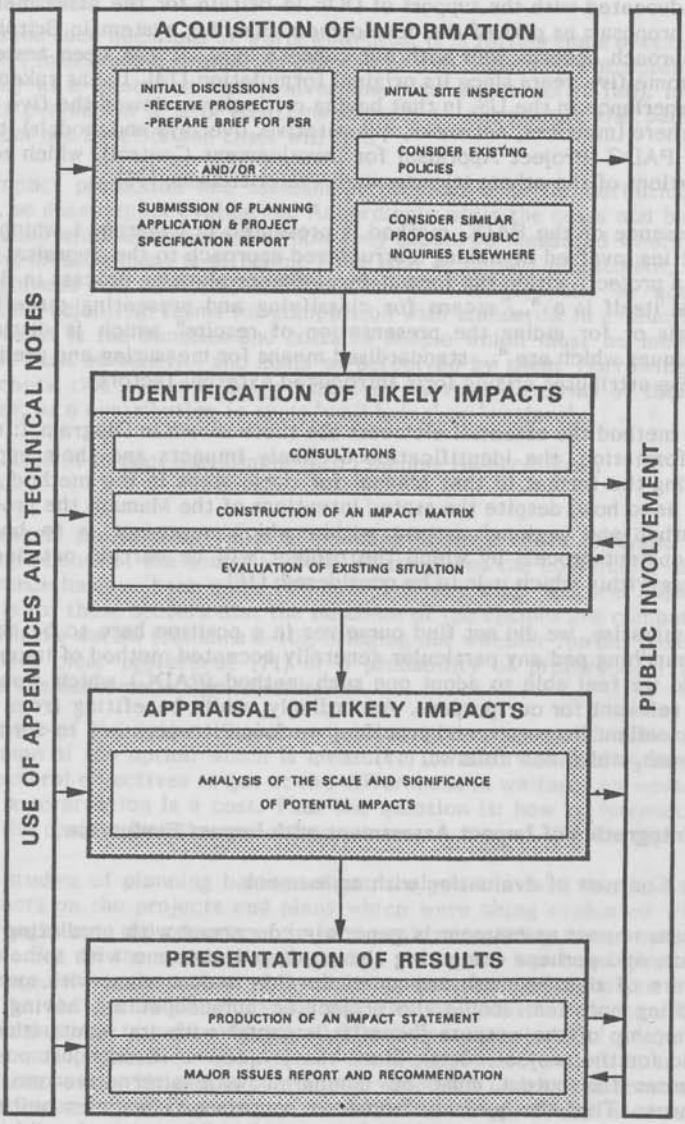
And it was here that the major impact of the US National Environment Protection Act, 1969, has been seen. Starting out with a preoccupation with the natural environment, the movement became involved also in the social and economic, which provided the opportunity for a launching pad from the natural resource to the socio economic impacts. In this it was the depth of scientific understanding of the natural environmentalists which has raised the levels of analysis directed to these other kinds of impacts. But this itself has become a drawback in that the scientific originators have tended to make over-technical and over-complicated the extension of the analysis from man's natural environment to his total environment, and in doing so have not been sufficiently concerned in producing conclusions from their analysis which are helpful in policy and decision making.

Thus the different streams are moving towards each other and aiding each other. But they are not yet in line, if only because there are different professional approaches and preoccupations. For example, the environmentalist has tended to be content with predicting the effects on the natural environment and the sociologists/anthropologists on human society, whereas the economist has gone on to the more difficult area of considering whether society would be better off or worse off by the use of scarce resources under the various welfare criteria, and introducing the basis for a judgement as a prelude to a decision.

4.2. Method of Impact Assessment

The simplest approach for our purpose would be to summarise in this section the principles and practice of impact assessment as a point of departure for the introduction of impact evaluation below (5.3). But for the reasons given above (4.1) this is just not practicable. While there are certainly common threads running throughout the principles and practice, and a common foundation laid in the pioneering programmes of the USA, there is considerable variety in the methods and techniques which are used.

But even when presenting the common thread of impact assessment, inadequacies do appear in terms of the requirements of the assessment process as a preliminary to evaluation. To demonstrate, reference is made to one particular presentation in a Manual of the impact assessment method to major development proposals. (13) This particular presentation is a useful one for the basis of this critique, for the following reasons.



Source: B.D. Clark **et al.** (1981)

Diagram 1. **Linked activities in the appraisal method**

It is advocated with the support of DOE in Britain for the assessment of development proposals as part of the development planning system in Britain, and thus in its approach accords well with our concern here. It has been tested in practice over some five years since its original formulation (14). It has taken advantage of the experience in the US, in that having critically reviewed the five main methods used there (matrices, networks, quantitative, overlays and models) it has devised a sixth, PADC (Project Appraisal for Development Control), which recognising the limitations of the others tries to meet the criticisms. (15)

The essence of the PADC method is presented in Diagram 1 which describes the activities involved in making a structured approach to the appraisal of the impact from a project, within the formal development planning process in Britain. The appraisal itself is a "... means for classifying and presenting material for impact analysis or for aiding the presentation of results" which is distinguished from techniques which are "... standardised means for measuring and predicting changes in these attributes arising form introduced external factors".

As to method the essential elements are those shown in Diagram 1: the acquisition of information, the identification of likely impacts and their appraisal. But in studying the format in that Manual for these steps in the method it is not at all clear as to how, despite the stated intentions of the Manual, the process relates to the urban and regional system within which a project is to be injected, the development process by which the project will be carried out and the planning process within which it is to be considered. (16)

To summarise, we did not find ourselves in a position here to be able to adopt as our launching pad any particular generally accepted method of impact assessment, nor do we feel able to adopt one such method (PADC) which would seem to be most relevant for our purpose. Accordingly, while benefiting from this illuminous and excellent literature and practice, we found it necessary to formulate our own approach, which now follows. (17)

4.3. Integration of Impact Assessment with Impact Evaluation

4.3.1. Contrast of evaluation with assessment

Whereas impact assessment is generally concerned with predicting and measuring impact, and perhaps comparing the predicted outcome with some standard (as a measure of significance), impact evaluation is concerned with answering a more searching question: should the project be pursued at all, having regard to the relationship of the outputs (benefits/impacts) with its inputs (the resources required for the project/costs). Since the project represents just one way of using resources the output must be compared with alternative uses of the input resources. Thus the question "should it be pursued" becomes quite searching. At the national level this could relate to all alternative possible uses of such resources in the economy: a daunting prospect. But even if the question be posed most narrowly, in the context of a decision having been taken to proceed with that particular project, then there is still the comparison needed of the most efficient way of carrying out the project, that is the best relationship of possible output with possible input. For this there must be a rigorous specification of the options which are in fact being compared, the "project options".

This comparative evaluation means that for every project there must be a datum (base) against which to compare it: what will the impacts be if the project is not carried out (the **without** situation) and what would it be if it were carried out (the

with situation) ?

In order to pursue such questions of worthwhileness, it is furthermore necessary to ask "to whom?". In economic evaluation this can be the individual entrepreneur or to the economy as a whole (on whomsoever the costs and benefits fall). In community impact evaluation it is in additionnecessary to disentangle the particular sectors on whom the benefits and costs will fall.

But just as impact prediction and assessment provides a better definition with measurement, so does impact evaluation. Accordingly while the costs and benefits should be included whether measured or not, they should be measured where practicable. But the measurement is different from that in impact assessment, where it is the magnitude and scale of the output whcih is generally in question, measured in some scientific terms for comparison with standards. In evaluation, as indicated above, it is the benefits and costs to people which must be measured. This being so it must be benefits and costs as perceived by **them.** Borrowing from economic concepts, this is seen as being perceived by them in terms of **their** sectoral objectives, as a contribution to their 'well being' or 'welfare'.

4.3.2. Incorporation of impact assessment into Planning Balance Sheet Analyses.

Diagram 2 presents the familiar table of planning balance sheet analysis, which sets out the ingredients of the analysis itself. It introduces the community sectors to whom reference has just been made (columns 1-3) and their sectoral objectives (column 6). It is for these sectors that the outcome of the options are compared by comparison with the datum (columns 8-10), bringing out whether for any particular sector there is a clear preference (11) or a probability of preference (12) and whether or not the outcome is significant (13).

In the comparison of the options (columns 8-10) it is the degree of achievement from the outcome of the option which is measured for each community sector in terms of its sectoral objectives to get at the differences in welfare; an advance is a benefit and a retardation is a cost. Thus the question is: how to introduce the impacts from the options into the evaluation.

In the earlier studies of planning balance sheet analysis the fact that it was the predicted impacts on the projects and plans which were being evaluated was not specifically brought out, the process of impact evaluation being somewhat within the 'black box'. But in the more recent studies the point has been given more specific attention, so much so that while planning balance sheet analysis is still considered to be the technique of evaluation the method itself is now better described as community impact evaluation. The 'impact evaluation' stems from the preceding considerations: and the term 'community' is to show that it is the **whole** array of impacts on the **whole** community which are under consideration and not simply particular impacts (economic, social, etc.) on **particular** sectors (e.g. transportation), or only those which are measured in money.

Thus it has been necessary to show how the assessment of impacts feeds into the evaluation. This we now proceed to demonstrate, by reference to Diagram 3.

Item	Community	No.	Project Variable	Impact Type	Sectoral Objective	Unit of Measurement	Comparison with Datum			Preference	Probability	Significance	Comment and Data
							1-D	2-D	3-D				
1	2	3	4	5	6	7	8	9	10	11	12	13	14
	PRODUCERS/OPERATORS												
1													
3													
5													
7													
.													
.													
.													
	CONSUMERS												
2													
4													
6													
8													
.													
.													
.													

Note: D = Datum.

Source: Any Planning Balance Sheet Case Study.

Diagram 2. **Planning Balance Sheet Analysis**

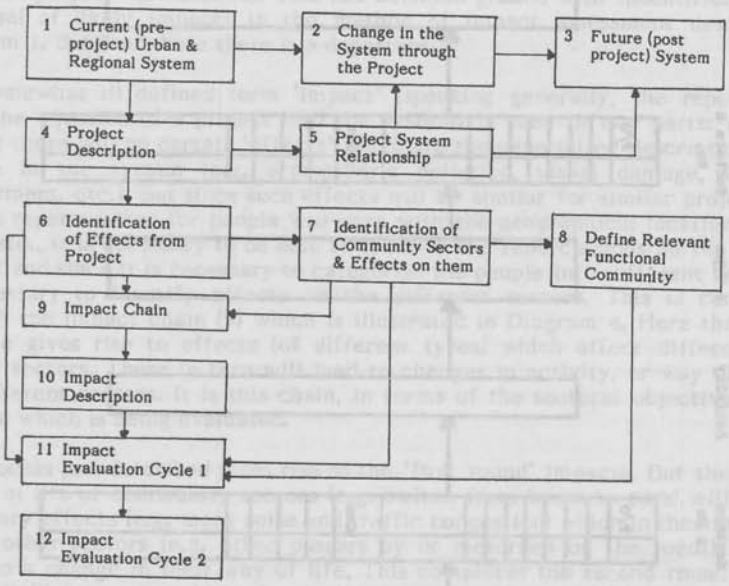


Diagram 3. **The steps leading to impact evaluation from a project**

In essence, the evaluation of a project is the evaluation of the change that the impacts of that project will make on the urban and regional system of which it forms part. This is expressed at the head of the Diagram (1-3) where the current system gives rise to the project which if implemented, would change the system. For this purpose we do not need to define the current and new system in total, but merely the change itself from the project.

This change injected by the project is amplified in the project description (4) which leads to comprehending the relationship of the project with the system itself (5). It is this change in the system which, as indicated, we need to incorporate into the project evaluation in Diagram 2.

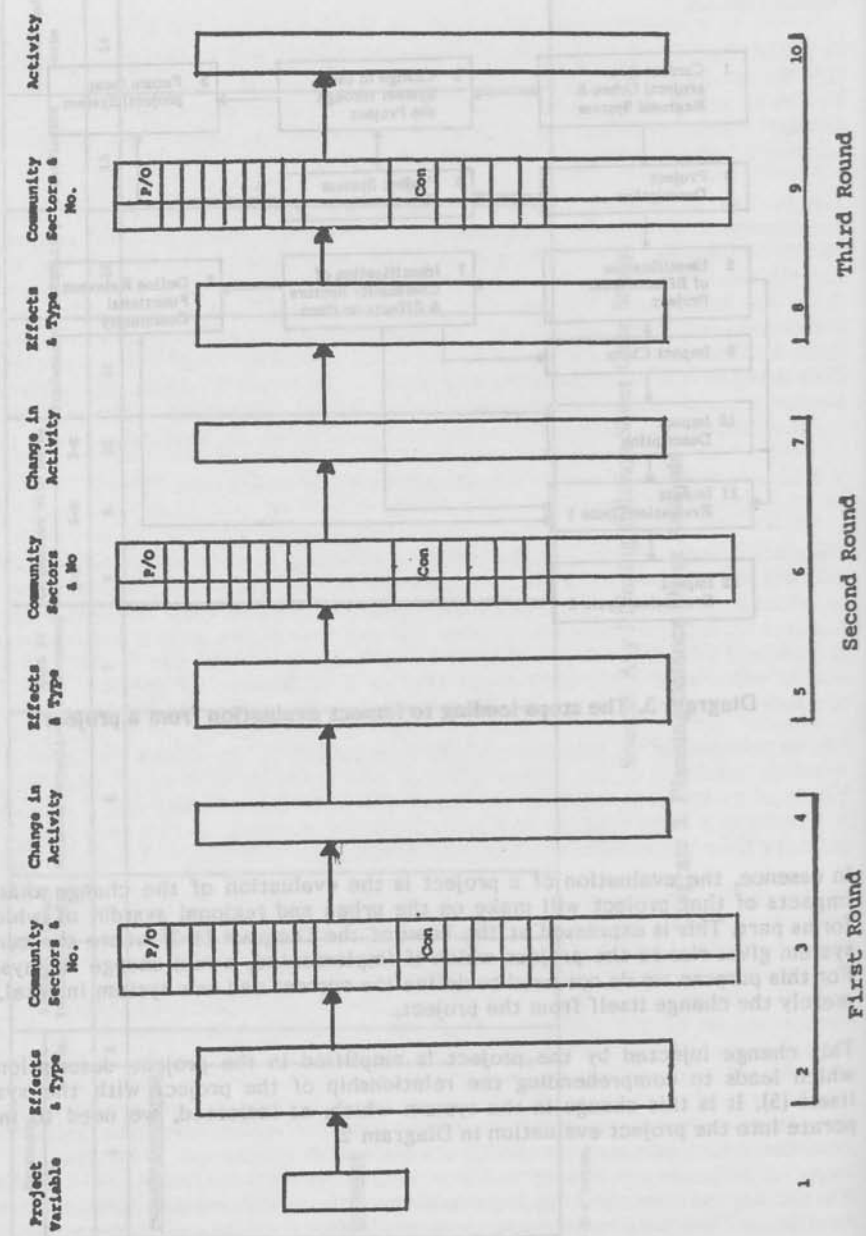


Diagram 4. **Impact Chain - Generalised Diagram for all effects**

The next critical step is the identification of the impacts which are likely to arise from the project variables (6). This has common ground with indentification and appraisal of likely impacts in the method of impact assessment described in Diagram 1. But from here there is a departure.

The somewhat ill defined term 'impact' (speaking generally, the repercussions from the injection of a project into the system) is seen in two parts: from the project there will be certain 'effects' which are the generalised description of the change in the system (e.g. atmospheric pollution, visual damage, increased employment, etc.). But since such effects will be similar for similar projects, and yet the repercussions for people will vary with the geographical location, size of town, etc., it is necessary to be able to identify the repercussions on the relevant people; and since it is necessary to categorise the people into different sectors, it is necessary to identify effects on the different sectors. This is carried out through the impact chain (9) which is illustrated in Diagram 4. Here the project variable gives rise to effects (of different types) which affect different community sectors. These in turn will lead to changes in activity, or way of life, of the different sectors. It is this chain, in terms of the sectoral objectives of the sectors, which is being evaluated.

The process just described gives rise to the 'first round' impacts. But the changes in way of life of community sectors (e.g. switch from buses to cars) will produce secondary effects (e.g. more noise and traffic congestion) which in themselves will affect other sectors (e.g. other passers by or motorists on the roads) and lead them to a change in their way of life. This completes the second round impacts which, in like terms, could lead on to third round and beyond, if it were possible to trace the repercussions.

Diagram 4 enables the critical step of impact evaluation to be taken (7), identifying the community sectors involved as a result of impact, leading to the ability to define the relevant functional community which is impacted (8) which is simply the aggregation of the different community sectors in (7). Clearly there will be difficulty in defining a tidy geographical boundary for such a community since it will be made up of a diversity of populations affected by a diversity of impacts.

Having identified the impact on the community sectors by first and second round, etc. (9) it is then possible to describe the impacts as they will arise from the project variables (6) through the impact chain (9) into the different community sectors (7). This impact prediction leans heavily on the impressive accumulation of principle, theory and practice in predicting and measuring impacts, which has grown up under impact assessment. The link therefore has been forged between the identification of the project variables (6) and the evaluation analysis in Diagram 2 (11).

A word remains about the distinction between Cycle 1 and Cycle 2 (11 and 12). Whereas in impact assessment the objective is to predict and measure the impacts, since the knowledge of measurement in various field is highly advanced we are faced with the familiar experience of impact measurement to a great degree of detail. But in impact evaluation the aim is to try and reach a conclusion on the comparative welfare of the community under the various projects, as an aid to the decision takers, within the constraints of time, etc. on their decision taking process. Accordingly the evaluation goes through an initial cycle which makes the analysis on the basis of readily available data, without committing time and

money to measurement beyond this point.

From this it could be a matter of good fortune that sufficient emerges from the Cycle 1 evaluation to give a clear indication of a preference for the decision maker; this has certainly been the experience on many planning balance sheet case studies. And even if the conclusion is not sufficiently clear to lead to a recommendation on the preference there will certainly have been learned a great deal about the options and their impacts (through the learning processes) leading perhaps to a rejection of certain options and the invitation to the analyst to use the conclusions to reformulate further options. But if the time has run out and the decision takers need to form a judgement, they will be better informed, so satisfying the criterion of evaluation: the ability to form a better judgement leading to decision than would otherwise have been possible.

But if the decision can be deferred, and it is thought that areas of uncertainty from an analysis without a full range of data can be reduced by further research, then the way is clear to proceed to Cycle 2 of the evaluation. Here the elements of the evaluation which require further clarification if a firmer recommendation is to be made are studied ad hoc, and the relevant data are collected with the hope that the conclusion and recommendation at the close of Cycle 2 will be sharper than that at the close of Cycle 1.

This in fact is the process which was carried out in one of the Greater Manchester Public Transport case studies, with Cycle 1 of the evaluation having been published and Cycle 2 currently underway. (18)

5. CONCLUSION

The main focus of this paper, the absorption of impact assessment into impact evaluation, is seen not as an isolated need but one which stems from the approach to evaluation presented at the outset, namely, to adjust our evaluation methods to a decision taker's criterion, bringing with it not a search for 'the method' but the need for greater synthesis in decision taking methods and techniques.

But the example chosen, impact assessment to impact evaluation, is symptomatic of the way in which principles, theory and practice of urban and regional planning have evolved. Since, by definition, such planning is both comprehensive in its coverage and relatively immature in its development, it is benefited throughout the years from the absorption of principles, theory and practice from academic professional disciplines outside the mainstream of urban and regional planning. There is only need to recall how the absorption of transportation, social and economic planning into the mainstream has considerably benefited urban and regional planning without losing its main purpose and thrust. And so it is with impact assessment. The importance of such impacts, be they social, economic, transportation or natural environment, has never been lost on planners. But it is the freedom and ability to embrace the important and wideranging methods and techniques of impact assessment which have grown up which has enriched the whole of urban and regional planning.

But by the same token, impact assessment must move towards planning evaluation. In practice, impact assessment has grown up in many countries independently of the planning field (notably in the USA where it started) and therby has introduced a distinct administration, practice and group of professionals outside the planning stream. Thus despite the weight of the work, in practice, it has not had its full potential bearing on the field of urban and regional planning and its

evaluation. Thus instead of being a contributory strength it has in many cases been a contributory weakness.

This is the danger to be avoided in the introduction of impact assessment, via the EEC into the European arena. (19) The introduction is healthy and will strengthen the concern for the environment. But unless it be incorporated and absorbed into the planning machinery and movement it could be counter-productive in effect. To some degree this is a matter of machinery, with each country under the EEC directive being responsible for how impact assessment is absorbed into its planning system. But while this may be so, a considerable advance can be made towards avoiding the disadvantages of overlap and conflict with those concerned with urban and regional planning, and in particular with policy, project and plan evaluation by absorbing the field of impact assessment into the field of evaluation.

It is hoped that this paper will make a contribution to that end.

REFERENCES

(1) Rates Act (1983)

(2) see chapter 1.

(3) Nathaniel Lichfield, Methods of Plan and Project Evaluation, in: Recent Developments in Planning Methodology, Proceedings of Netherlands Colloquium, October 1977, Ministry of Housing and Physical Planning, The Hague, (1978), p. 150.

(4) This is an amplification of Lichfield, see ref. (3), pp. 154-158.

(5) Nathaniel Lichfield & Partners, Achieving Value for Money in Public Transport: Approach and Methodology (Greater Manchester Passenger Transport Executive, 1983).

(6) Andreas Faludi, The Rapporteurs Summary, Method of Plan and Project Evaluation, see ref. (3), p. 161.

(7) Nathaniel Lichfield, Peter Kettle and Michael Whitbread, **Evaluation in the Planning Process**, Pergamon Press, London (1975).

(8) see chapter 1.

(9) Nathaniel Lichfield et alia, 1975, p. 96.

(10) Nathaniel Lichfield, Evaluation Methodology in Urban and Regional Plans: A Review, **Regional Studies** (1970), Vol. 4.

(11) B.D. Clark, R. Bissett & P. Wathern, **Environmental Impact Assessment: A Bibliography with Abstracts**, Mansell (1980). And see for a review of impact studies of various kinds, J. McEvoy and T. Dietz (eds.), **Handbook for Environmental Planning: The Social Consequences of Environmental Change**, Wiley, London (1977).

(12) Norman J. Glickman (ed.), **The Urban Impacts of Federal Policies**, The Johns Hopkins University Press, Baltimore (1978) and N. Lichfield (ed.), **The Urban Impacts of National Policies**, OECD, Paris (1983).

(13) B.D. Clark, J.K. Chapman, R. Bissett, P. Wathern & M. Barrett, **A Manual for the Assessment of Major Development Proposals**, HMSO, London (1981).

(14) B.D. Clark, J.K. Chapman, R. Bissett, P. Walthern, **Assessment of Major Industrial Applications: A Manual**, HMSO, London (1976).

(15) B.D. Clark, J.K. Chapman, R. Bissett, P. Walthern, **Environmental Impact Assessment in the USA: A Criterical Review**, HMSO, London (1978).

(16) B.D. Clark, et al., see ref. (13), Ch. 3 and Appendices A & B.

(17) cf. Nathaniel Lichfield and Uri Maninov, Land Use Planning and Environmental Protection: Convergence or Divergence, **Environment and Planning**, Vol. 9 (1977)

(18) Nathaniel Lichfield & Partners, Achieving Value for Money in Public Transport: Trafford Park Case Study, Vols. 1, 2, 3 (Greater Manchester Passenger Transport Executive, 1983).

(19) Draft Directive Concerning the Assessment of the Environmental Effects of Certain Public Projects, Council Reference 7972/80; Commission reference com. (80) 313 final. Official Journal Nr. C169 (9 July 1980) p. 4.

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EXPLICIT EVALUATION IN A PERIOD OF DECLINE AND SCARCITY

J. Buit
Dept. of Urban and Regional Planning
Free University Amsterdam

1. EXPERIENCES WITH DUTCH EVALUATION STUDIES

It is a difficult and hazardous task to discuss the possibilities and the desirabilities of an explicit evaluation of complex, spatial, planning problems in a period of decline and scarcity. Because of the many uncertainties arising from a future of decline and scarcity, and because my knowledge of the applications of evaluation of spatial planning principles and strategies, is by and large, limited to the Dutch spatial planning scene, there is only one thing certain beforehand. My ideas on this topic are relatively uncertain, not well cristallized, and are by choice, sometimes rather extreme and exaggerated. What I am trying to accomplish with this paper is to provide a starting point for further study and discussion.

The following points are covered. First, the application during the last ten to fifteen years of evaluation from both simple and complex spatial problems in the Netherlands is discussed. What were the experiences with evaluation in a period of growth and abundance? On what scale, and for which types of planning problems was evaluation used? Looking at the contents, what are the significant traits of evaluation, and how do these compare with the dominant characteristics of the evaluation methods as designed by Lichfield and Hill, the patriarchs of evaluation?

Which factors have been important in a period of growth and abundance in the use, and non-use, of evaluation in the process of preparing and implementing central spatial decisions? Such an analysis of the experiences with evaluation during recent years is indispensable, because it can give essential indications of what we can expect of the application and readjustment of evaluation during a period of decline and scarcity. After the presentation of an outline of the track record of evaluation in a period of growth in the Netherlands, out attention will shift to the expectations of evaluation in a coming period of decline and scarcity. First, I will give a synopsis of the most urgently needed readjustments in the contents of evaluation, readjustments needed to enhance its obility to serve as an indispensable vehicle or guide for answering the types of questions on spatial planning which, as a consequence of decline and scarcity, will demand central attention from politics and society.

The second point I wish to deal with is a prediction on the chances for successful, explicit evaluation of complex spatial planning proposals and implementations in a situation in which government and society are struggling with decline and scarcity.

If we look at the applications of evaluation in spatial planning in the Netherlands during the last one or two decades, then at first sight there is some cause for satisfaction. There is a large and rapidly growing body of publications which compare spatial options, using evaluation methods either on a firm quantitative

base, or in a more impressionistic and qualitative manner such as the scenario method (1).

Many planning reports on local, regional, and national spatial planning tried to select the best option with the aid of evaluation methods. Moreover, there are many studies of a type in which a comparison is made between one future spatial situation and the existing spatial situation, also using explicit evaluation. In the application of evaluation, a broad spectrum of methods is used: there are studies of the cost-benefit type concentrating on market potentials; studies of the planning balance sheet type concentrating on market potentials and consumer preferences; and studies of the goal-achievement matrix type concentrating especially on goals within the government sphere. It is striking that the overwhelming majority of evaluation studies is of the ex-ante type. Ex-post evaluation of realized spatial proposals is scarce (2). Finally, there is a rather small quantity of publications with an explicit and systematic evaluation of the expected or real results of existing, proposed, or abolished spatially relevant instruments, such as infrastructural works, subsidies, differentiation of housing types, selection principles for assigning new or existing houses to people, and so on. These evaluations are based on central goals and objectives.

Another point worth mentioning, as it has led to a satisfactory application of evaluation, is that after a period of trying out and learning evaluation, something like a standard evaluation procedure with a firm and regular repeated methodology, and the use of regular criteria for spatial decisions, is becoming visible in several government sectors. Situations in which these are applied are: the selection of new, large housing estates, the tracing of new autoroutes, the building of new hospitals, the realization of new shopping centers, the opening or closing down of schools, and -most recently- the siting of projects with negative effects for our natural/physical environment (3). Looking at this list there is one essential conclusion to be drawn; successful applications of evaluation in the last decade are found especially in situations where decisions had to be made regarding new construction and expansion plans, and in projects and plans where primarily the intra-sectoral comparison of alternatives, on the basis of intra-sectoral goals and limiting conditions, takes place.

It is quite clear that the degree of success in the application of evaluation in the past years has a lot to do with those objects that are characteristic for a period of growth and expansion. In a period of growth and expansion, evaluation is rather successful because it is used for comparing alternatives on new building projects with relatively few negative effects for the interested groups. The situation is quite different for renewal projects. Because in the case of new building projects it is used for comparing spatial alternatives within relatively autonomous planning sectors with highly consistent, instead of conflicting, internal packets of objectives, there is no necessity to take into account the effects of the decisions outside the sector in question. In a period of growth and expansion, integrated spatial planning can be described as, roughly speaking, the sum of the sectoral spatial plans, each the result of optimization-oriented evaluation within the sector itself.

This first impression of a strong relationship between the degree of use and the more specific content of evaluation, and the type of spatial planning questions and problems which are naturally central in a period of growth and expansion, becomes much stronger if one looks carefully at the dominant features of the evaluation matrices used, and at the spatial questions and options for which real evaluation exercises have been totally absent.

Regarding the significant characteristics of the contents of the evaluation matrices, the following points can be made:

1. There is a strong tendency to use, where possible, a set of non-conflicting objectives in public evaluation studies (especially within sectors). Also, in inter-sectoral fields, there is an attempt to eliminate conflicting objectives in order to promote a policy of integration and coordination. This means that evaluation studies, as a consequence of this deliberate strategy within the governmental sphere in the selection of objectives, are less pluralistic than one would expect in a pluralistic and conflict ridden society. Even in a period of growth and expansion, there is much reluctance to show, via evaluation studies, problematic and conflict-raising aspects of spatial choices (4).

2. The previous conclusion seems even more correct if one looks at the specification of effects, per interest group, in the evaluation matrices. Contrary to the methodological instructions of Lichfield and Hill, one never finds an evaluation study in which the effects of alternative, spatial options are specified for a broad variety of interest groups. All questions on the equity-efficiency topic are systematically avoided. In a policy context, it is much less dangerous, even in times of growth and expansion, to use rather innocent and general objectives in an evaluation matrix, (objectives such as accessibility, flexibility, and livability) rather than to show financial and non-financial costs and benefits for each interest group.

3. Ex-post evaluation studies, as mentioned before, are scarce. The results, when showing heavy positive and negative consequences of spatial strategies in the light of conflicting public objectives, are, in essence, not very welcome in the governmental planning offices. They show all too painfully the conflicting aspects of selected spatial strategies and the fact that the choice of these strategies, in many cases, is based more on politics and the influence of interest groups than on knowledge of expected or real societal consequences. Evaluation studies are more accepted if they have to do with matters that are not all too controversial for both government and society.

4. All evaluation studies compare alternatives which supposedly are already realized. This means that all aspects of realization and implementation (degree of societal acceptance, pace of approval, political problems, time and costs of delays in realization, etc.), which in fact have such an enormous influence on the final selection of a specific project, are kept out of the traditional evaluation exercise. This is one of the main reasons why the selected alternative is frequently not the one recommended in the evaluation study, or why none of the evaluated alternatives is ever realized (5). This weakness of evaluation has to do with the fact that the political field long considered explicit attention for the implementation side of evaluation as less desirable: it only serves to draw attention to the project beforehand, reducing the room for negotiations, etc.

The recent boom in strategic choice exercises and uncertainty analysis (vehicles for scientific emancipation of procedural planning just as evaluation formerly served as an instrument for scientific emancipation of substantial planning) has hardly altered this situation in my opinion. The use of an isolated, strategic choice tends to produce the situation where difficult spatial decisions, that means decisions with uncertain relations to other decisions, must be postponed until more certainty is obtained. But there is no effort to state the consequences of such a delay in terms of extra costs, continuing housing shortages, societal opposition etc., in relation to the results of substantive evaluation. So there is no explicit and real integrated evaluation exercise in which the evaluation matrix shows both

what substantive gains and losses can be expected for each alternative, and what procedural gains and losses can be expected for each alternative. Only then is a decision on the spatial options possible, a decision in which decisionmakers can weigh substantive and procedural pros and cons: do they wish a solution with relatively good societal results after realization but with many difficulties and delay before realization, or one with less good results after realization but with a smooth and fast realization?

5. There are some indications that, in politically more complicated and controversial evaluation studies, there is a tendency not to mention alternatives which are seen in the political field as less desirable, or to select evaluation criteria so as to be certain that the politically desirable alternative will gain the most positive points (6). In such situations there is a would-be evaluation, and one is reminded of Bachrach and Baratz who state that there is always an attempt in politics to exclude all too controversial topics from the public agenda (7).

6. In most evaluation studies, there is an effort to optimize in the light of positively valued objectives on the benefit side, reflecting in this respect a society of growth and expansion. Only recently are there signs in evaluation exercises that aspects such as cost reduction, use of existing overcapacity in infrastructure, services etc. and satisficing in a situation of scarcity and problem-diminishing, are getting more attention (8).

7. Almost no evaluation studies showed any interest in future operating costs for government and society, in permanent costs for mobility, energy, etc. for users after realization of the preferred plan, and for side-effects outside the territory of the plans (9). All these traits are logical in a period of abundance and growth, when such knowledge is not very essential in the weighing of alternatives.
The overall conclusion is that the contents of evaluation studies in the past ten to fifteen years show several, typical characteristics, which can only be understood to be a result of influences coming from a governmental machinery which strived for rather non-problematic decision-making, and from a society in which, thanks to growth and abundance, evaluation was frequently a non-problematic and non-threatening matter.

This tendency in the government sphere to use evaluation only in non- (or marginally) problematical spatial decisions, and to neglect and avoid it in more highly problematical spatial decisions (such as situations in which political urgencies, more than rational factual investigations, are crucial for decision-making), becomes visible when one looks at the list of spatial decisions taken without serious explicit and published ex-ante, and until very recently ex-post, evaluation studies.

No beforehand evaluation study was made for such central, spatial decisions as: the number and size of new towns and their distance to central cities; the selection, number, and quality of growth centers in the more peripheral parts of the country; the removal of national, governmental offices from the congested Randstad; the founding of universities in peripheral regions of the country; the urban renewal policy in terms of desirable spatial and population characteristics after the renewal operation; the degree of differentiation of urban environments within cities; the segregation or mixture of urban population groups; the proposed creation of the monstrous Noron-new town, half oriented to Haarlem, half to Hoofddorp; the permitted population growth of 1% per year or the equivalent of the natural population growth for all regions in the countryside incl. the green heart of the Randstad, and, finally, for the urban traffic circulation plans. It is no accident that precisely these topics are without published evaluation studies: they

are all spatial decisions taken in the light of highly conflicting political and societal views, with many unproven and exaggerated claims, with great uncertainty over the types and quantity of subsidy regulations, and so on. For the same reasons, systematic evaluations of instruments in terms of costs and results are also scarce.

To give only one example: it is only a few months ago that the first, rough study appeared comparing the effectiveness of all spatial and non-spatial instruments used to promote public transport, in spite of the fact that the promotion of public transport has, for years, been judged by the government to be of central importance (10).

The scarcity of good ex-ante and, even more, of ex-post evaluation studies on so many central aspects of spatial policy in our country, even in a period of growth and expansion, must be partly seen as the outcome of political decision making. Evaluation of politically preferred spatial options is too risky, because it provokes discussion and conflicts, and because decisions on political grounds can be forced through more easily without exact knowledge of the consequences of such decisions.

But there are undoubtedly other causes as well. In a period of growth, there is less chance for real, negative effects elsewhere in the spatial system due to building projects; at the most, there is a differential rate in positive growth effects. This means that there is only a moderate interest in such evaluation exercises. To mention just one example: as long as new town realization had no serious negative effects for the growing central cities, there was no need to evaluate such effects.

The relatively small chance of serious negative effects due to building projects, effects which manifest themselves elsewhere in the spatial system and in other sectors of land use during a period of growth, makes it understandable that there is a emphasis in regular evaluation on intra-sectoral, and not extra-sectoral, consequences of spatial options for each sector. There is no urgent need to evaluate consequences in other sectors, such as the consequences of hospital building for the housing environment, of inner city traffic-circulation plans not only for the traffic, but also for businesses and living conditions outside the inner city, etc.

One final and wellknown point can be mentioned as a cause for scarcity of evaluation studies of the ex-post type. Such evaluations must follow effects and processes for longer periods of time, before and after realization of a spatial option or between situations with and without the realization of a spatial option. This means that the results are too late for further decisions, especially if there is, in the meantime, a substantial change in objectives. Moreover, the results are not always exact and reliable, because it is very difficult to isolate direct, and especially secondary effects of the realization of a spatial option from the effects of all the other factors causing change and development in the spatial system (11).

All together this brief outline of experiences with evaluation in the Netherlands during the last fifteen years leads to the conclusion that some satisfaction is justified, but there is also cause for dissatisfaction. On the one hand there are good, quantitative and qualitative results, but on the other hand, there is a disappointing scarcity of application, and disappointingly meager contents of such applications when dealing with 'hot' spatial topics in a period of growth and expansion. Valuable insight can be gained, however, in the desirability, but also in the possibility of application of evaluation in the coming years of scarcity and decline.

2. RECOMMENDATIONS ON THE CONTENT OF EVALUATION STUDIES IN A TIME OF SCARCITY

Trying to formulate recommendations for relevant changes in the art and contents of evaluation of spatial questions so as to promote the possibility of fruitful application, and answering urgent political and societal questions in time of scarcity, means that one must have ideas on such questions. I shall therefore try to give some personal and subjective impressions of the consequences of scarcity, both for the types of questions dealt with as well as for the potential role of a changed evalation content in answering such questions. I would like to make the following points.

1. In a period of extreme scarcity of (risk-bearing) money, there will be an urgent need for a much more complete financial evaluation of alternatives on spatial strategies and projects. In such a financial evaluation, it will be necessary to include all financial aspects during the period of preparation (incl. costs of bargaining and bargaining time, delays in preparation, etc.), during the period of realization (incl. delays in realization), and during the whole lifespan of the project, or a strategy after realization (incl. risks and costs of vacancy). Moreover, it is of great importance to look at costs and benefits not only for the government and/or investors, but also at the financial consequences of the plan for future users.

2. In a period of scarcity of mony for both consumers and commercial and non-commercial activities, there will be a need for evaluation of spatial alternatives which maximize cost-reducing aspects such as home-work distances, more intensive use of existing buildings, and infrastructure inside and outside the plan area, and so on.

3. In a period of scarcity of employment, it also will be of great importance to concentrate on employment effects of spatial alternatives: how much and what type of employment, especially for women, will be gained or lost directly or indirectly as a consequence of spatial alternatives and/or instrumental alternatives?

4. In a society with scarce and threatened resources, there is renewed interest in efficiency and redistributional effects of spatial alternatives. This means that evaluation must pay more attention to equity-efficiency aspects of alternatives, and to the effects for each relevant interest group in society (12).

5. In a period of scarcity, there will not be a high rate of construction activity. This means that the main focus of attention in evaluation must shift from planning building projects to strategies and instruments to realize a better (that is cost-reducing and/or income-enlarging) locational pattern of activities within a nearly fixed stock of existing buildings.

6. Scarcity will lead to a call for deregulation, for less and less complicated instruments with a high degree of effectiveness in terms of goals/objectives and costs. Evaluation of existing and proposed instruments will be a relatively new and promising field of regularly repeated evaluation research.

7. Scarcity will lead to more intense conflicts between government and interest groups, and between the various interest groups themselves, on real or imagined consequences of real or proposed spatial plans and strategies. There will be in such a situation, much demand for counterexpertise evaluation using the same or other evaluation criteria to confirm or correct evaluation results of certain planning authorities or other interest groups.

8. A period of scarcity for the government will lead to a reduction of public goal formulation as the basis for spatial plan evaluation. In the coming years, evaluation must pay more attention to consumer and producer preferences, and to increasing the options for consumers and producers in the comparison of spatial plans and strategies.
9. In a period of scarcity, there will be more intense, and sometimes unexpected, reactions before, during, and after realization of spatial plans and strategies. It will be indispensable for evaluation in the coming years to give full attention to second and third round effects of alternative plans and strategies within, and especially outside, the planned area. If ignored, one would be closing one's eyes to the sad fact that second/third round effects are, in many cases, plainly contrary to first round effects (13).

10. In a period of scarcity and financial uncertainty, the risk of abandonment or altered realization of spatial plans is very high. This means that the evaluation exercise, as part of contingency planning, must give information beforehand on the consequences of changed realization, or provide the option of quickly obtaining this additional information as soon as an altered realization becomes necessary.

11. In a period of scarcity, there will be a need for re-evaluation of the more costly parts of existing spatial policy. It is plausible that the new town policy (as opposed to the compact central city philosophy), the growth cities policy, the green heart policy, and the urban renewal policy will demand intensive evaluation studies in the next few years.

12. Scarcity will lead to lowered chances for realization, especially if financial problems go hand in hand with expected realization problems. In integrated evaluation, there must be emphasis given to all those aspects of realization which make it either easier or more difficult to carry a project out, such as procedures, instruments, subsidies, action groups, societal and political resistance, chances for commitments and contracts between interest groups, and so on. Evaluation without explicit exploration of this implementation side will be, in fact, without much value in a period of scarcity.

All these adjustments, in my opinion, are important in ensuring that evaluation will be applied in answering questions on the selection and implementation of spatial strategies. Especially the last point, I think, is of utmost importance: only an explicit and integrated evaluation of procedural and substantive aspects of alternative plans and strategies will help to answer the most urgent option-questions in a period of scarcity. Only then will it be clear what the substantive and procedural effects are of options which favour either the contents or the implementation of the plan. In a time of scarcity, procedural possibilities will, without doubt, play a greater role in choosing options than the substantive aspects of the option.

3. CHANCES FOR EVALUATION STUDIES IN A PERIOD OF SCARCITY

After this outline of the necessary adjustments of evaluation in times of scarcity, it is now time to address the question of the chances for the use of explicit and integrated evaluation studies for answering complicated questions from both government and society on spatial choices and instruments. Is there a large and growing field of work to be expected for evaluation in such a time, or will there be much less demand for evaluation in comparison with the recent years of growth and expansion?

It cannot be denied that in many respects, the chances for explicit and integrated evaluation research in the coming years will be smaller than in the years of growth and expansion. It is enough to point to the following facts:

1. The deregulation of spatial instruments, especially of financial instruments, will result in a diminished need for evaluation, for a smaller number of instruments.

2. In a period of scarcity, and thus of growing societal and political conflicts over the distributional effects of spatial plans and strategies, there will be great reluctance from the side of the government to provide opportunities for conflict-raising, explicit, evaluation studies, especially if the evaluation traces effects per interest group.

3. In a period of scarcity, realization of projects will be very difficult, and difficult negotiations can be expected before commitments and/or contracts can be made. In such a situation, many participants in the preparation and realization of the project will be against explicit, published evaluation of the project, both ex-ante and ex-post, because it endangers the commitments and contracts already made if one or more participants in the negotiations -using the evaluation results- ask for a readjustment of the commitments or contracts to ameliorate their position.

4. In a period of scarcity, there will be relatively few, large, construction projects, and it is exactly to this type of expansion plans that evaluation was frequently applied in recent years.

5. In a period of scarcity, there will be less room for an extensive set of goals and objectives from the side of the government. Necessarily concentrating actions on problem amelioration and ad hoc incrementalism will provide less room for extensive evaluation studies using extensive sets of goals and objectives.

6. In a period of scarcity, there is a tendency to decentralize governmental instruments and decisions to lower or local levels. On such levels, decisions are relatively simpler, and the necessity for evaluation studies as a basis for decisions is much less.

7. In a period of scarcity and uncertainty about the future, there will be less political will to speak about and to take decisions on vital and unpopular spatial options such as enforced mobility of people and of businesses. Evaluation on these matters will be postponed or canceled.

8. In a period of scarcity, evaluation research will be judged rather negatively, because such research is often costly, and the results are frequently too late to influence urgent decisions. Decisionmakers will, in such situations, mercilessly cancel intended evaluation studies.

Altogether, it looks as if there are fewer possibilities for evaluation studies in a period of scarcity then in a period of growth and expansion, despite the fact that, as was argued earlier, the quantity and quality of evaluation studies was, even then, quite modest. On the other hand, there are also signs and forces which point in the opposite direction, the direction of a greater demand for evaluation studies in a period of scarcity, as compared to a period of growth and expansion. It is thus plausible that:

1. In a period of scarcity and deregulation, the remaining instruments will require

intensive and repeated evaluations based on their (re-)distributional effects and effectivity.

2. In a period of scarcity, there are great dangers of unexpected side, second, and third round effects of spatial strategies and plans. Investigation and evaluation of these effects will have a high priority.

3. In a period of scarcity every interest group will, in the name of survival and subsistence, want to watch carefully the real and expected effects of spatial plans and strategies. Interest groups themselves will do a lot of evaluation research, or they will rightfully claim that the government must sponsor this research.

4. New ideas on the necessity of more or less radically changed spatial strategies (for example compact central cities, free enterprise zones, reduction of distances, more intensive use of underutilized buildings and infrastructure) demand testing via evaluation research.

5. In a period of scarcity, the hard core of spatial decision-making will change from choosing between alternative locations for town expansions to the choice of better locations for activities within the existing stock of buildings. This means more a change in urgent evaluation research, and not so much a reduction.

6. In a period of scarcity of labour, there will be a new and strong interest in the effects on existing employment, and in the employment and economic potential of spatial strategies and plans. This provides new opportunities for evaluation research.

7. In a period of large risks and uncertainties for fruitful realization of spatial plans and strategies, evaluation of realization aspects and chances, including procedural impediments, will be of utmost importance for decisionmakers, even if they are reluctant or unwilling to publish the results of such evaluation research.

8. In a period of scarcity and conflict, and with a government that will try to avoid evaluation studies on tricky and problematic spatial choices, science and society must pose the difficult and critical questions on the effects of such politically based spatial choices. The fact that it is impossible to answer these questions without a thorough evaluation research will help to ensure that such evaluation research will be carried out.

The end-result is difficult to predict. Some factors enhance, others restrict the possibilities for evaluation research on spatial strategies in a period of scarcity. Although the outcome is unsure, one thing is certain: in a situation where factors work against each other forming an unstable balance, not much is needed to push the outcome in one direction or another. Only a small amount of action or influence might be sufficient to produce the decision to apply full-scale, evaluation research on spatial questions as an indispensable instrument to answer such questions. Science and society, in this situation, can effectively influence the amount of evaluation on spatial questions. They can be of influence by making clear to politicians how urgent such research is. Science, in this field of urban and regional planning, can also be of influence by publishing evaluation applications which are adopted to and fruitful for answering the types of policy questions which typically arise during times of scarcity.

NOTES

(1) To give one example of a report based on the scenario method: see "Three scenarios for the preparation of the Urbanization Report", Studierapport Rijksplanologische Dienst, The Hague (1975)

(2) Ex-post evaluation studies can be found (although also very seldom) in urban traffic planning and in shopping studies.

(3) With respect to new large housing estates mention can be made of reports from the Provinciale Planologische Diensten ("Dutch Provincial Planning Authorities") of North and South Holland and recent reports from the planning departments of the city of Amsterdam and Utrecht. For the allignment of new motorways see the contribution of Van Staalduine (this volume). For planning of hospitals see the regional reports of the College voor Ziekenhuisvoorzieningen. For new shopping centers see the many reports of the Centraal Instituut voor het Midden- en Kleinbedrijf. For schools see reports from the Ministry of Education and, finally, for environmental impacts there is the recent decision to prepare environmental impact statements.

(4) Within sectoral fields one can point, for example, to evaluation exercises for hospital planning in which only hospital-oriented objectives (viz. no urban planning objectives) are used. Other examples are the fight for objectives pro- or contra- the Markerwaard (see also the contribution of Sorber in this volume) and the lack of economic objectives (costs and benefits for enterprises and population) in evaluation reports on new building sites. The gravel extraction evaluation study of Bennema c.s. (this volume) is also interesting in this respect: non-governmental interest groups were not allowed to participate in the study in order to create a relatively 'quiet environment', but even now there are difficulties about yes or no publication of the final report of this study.

(5) It is one of the major reasons, for example, why the Noron-new town will only be partially realized (excl. Vijfhuizen-Noord), why Havens-Oost as a new expansion of Amsterdam will not be realized (conflict Amsterdam pro, central government against realization) etc. Even in a recent Amsterdam report on future building sites only one of the eight central objectives dealt with the implementation time (Report "Potentiële Nieuwbouwlocaties", Dienst Ruimtelijke Ordening Amsterdam).

(6) One older example is the report "Globale Visie", in which four highly comparable building site plans for Midden-Utrecht were presented, but in which no attention was paid to the alternative of one (or two) real big new town(s). More recent is the discussion (and 'fight' for proper objectives and criteria) around the 'new towns' versus 'compact city' dilemma.

(7) P. Bachrach and N.M. Baratz, "Power and Poverty; Theory and Practice", Oxford University Press, New York (1970)

(8) In a more descriptive way than in the form of a systematic evaluation matrix one may find these aspects in the Dutch National 'Structuurschets Stedelijke Gebieden', Staatsuitgeverij, The Hague (1983)

(9) User costs are not distinguished in the recent evaluation exercise of the Netherlands Economic Institute, in which the costs of a housing project in a new town are compared to an expansion plan of a central city; nor areuser costs presented in a study of the Stichting Economisch Onderzoek about the costs for the government of building in a renewal area versus building outside the town.

(10) Mentioned by ir. G.H.A. Hoogenboom, head of the research department of the Ministry of Traffic in a recent lecture on this topic.

(11) One of the problems of secondary effects is the fact that such effects are the result of active reactions in terms of spatial decisions of firms and people after a period of passive ac-

ceptation of direct effects from spatial plans or strategies. An example of such a study on rounds of ex-post effects and problems of investigation is: J. Buit, "Repercussions of the Inner-City Plan of Groningen for Enterprises and Inner City Environment" (2 reports), Amsterdam, Free University, Dept. of Urban Planning (1979, 1981).

(12) Trying to find relevant criteria for questions related to equity (a difficult matter as is explained by Miller in this volume) must be seen as very essential.

(13) There are signs especially on second and third round migration of enterprises and people (chains of relocation), that the effects of these migrations in terms of governmental objectives are far less favourable or even pure negative in comparison with first round migrations. It has partly to do with the fact that second/third round relocations of firms and people are far less controllable in terms of goals and objectives.

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EVALUATION AND INFORMATION: A SPATIAL PERSPECTIVE

P. Nijkamp
Dept. of Economics, Free University Amsterdam

INTRODUCTION

In the post-war period, a wave of information has flooded many societies, east and west. Both private and public agencies have increasingly become reliant on information systems as indispensable tools for planning and decisionmaking. This information explosion has been induced by the introduction of computers, microelectronic equipment and telecommunications services. These advances in modern technology have offered an enormous potential by using information in a logical and well structured way for handling complex problems of choice and decision (see also Burch et al., 1979, and Debons and Larson, 1983).

The information explosion is due to various causes: the need for proper insight into complex societal processes, the high costs incurred by wrong decisions in modern agencies, conflicting interests among decision agencies ("information is power"), the rise in data collection equipments in many statistical offices, the progress in statistical and econometric methods, and the emergence of modern computer hardware and software possibilities (computer consulting systems, e.g.).

In the next section a number of issues with respect to evaluation and information will be discussed. It will be outlined that data for evaluation should be represented in an operational form in order to make the actual choice issues as transparent as possible. This implies some form of information management, which is additionally discussed in more detail. The paper will be concluded with some general remarks on flaws and prospects of spatially-oriented (regional) information systems.

EVALUATION AND INFORMATION

Evaluation aims at rationalizing planning and decision problems by systematically structuring all relevant aspects of policy choices (for instance, the assessment of impacts of alternative choice possibilities). Evaluation is usually not a one-shot activity, but takes place in all phases of decision-making (for instance, on the basis of learning principles). In addition, a systematic support to complex planning and decision problems presupposes a balanced treatment of too many details and too little information. Besides, the results of an evaluation procedure have to be transferred to policy-makers in a manageable and communicable form, particularly because the items of an evaluation problem are usually multidimensional in nature (including incommensurable or even intangible aspects). Finally, it has to be realized that the "planning environment" is usually highly dynamic, so that judgements regarding the political relevance of items, alternatives or impacts may exhibit sudden changes, hence requiring a policy analysis to be flexible and adaptive in nature. Rigid evaluation techniques run the risk that an evaluation does not cover all planning issues in a satisfactory way.

Any evaluation requires appropriate information. The aims of the evaluation however, may be different and depend on actual institutional and administrative interest. Three broad categories of behavioural paradigms may be distinguished for public decision-making:

- "optimizing" behaviour
- "satisficing" behaviour
- "justificing" behaviour.

Although the majority of formal evaluation techniques is focusing attention on the first category and to a lesser extent on the second category, in policy practice evaluation is often used as a means of justifying policy decisions, even if the actual decisions are not in agreement with "optimizing" or "satisficing" principles. In any case, however, relevant data for a policy judgement have to be collected. Such data should be represented in an operational form in order to make the actual choice issues as transparent as possible.

Any policy decision will affect the welfare position of individuals, regions or groups in a different way. Consequently, the public support for a certain policy decision will very much depend on the distributional effects of such a decision. Thus, in general, it is advisable to design or use evaluation methods that try to assess the pros and cons of a certain choice alternative for separate groups or regions. Information on such gains and losses are not always cardinal in nature, but also qualitative, fuzzy or verbal information may provide a meaningful input for a policy analysis. Altogether, spatial and/or social referencing of information is highly desirable to make evaluation more effective.

Beside the (institutional or administrative) structure of a decision problem, the specific plan evaluation method to be used will also determine which data are requested for the policy analysis at hand. For instance, checklist approaches, cost-benefit studies, planning balance sheet techniques, goals achievement methods, multiple criteria analyses, multiple objective programming models have all their own specific data needs. Clearly, assuming a cyclical model of planning implies also that the relevance of an evaluation technique has to be judged in light of the available data. In this regard, it is worth noting that monitoring is a necessary ingredient of an adaptive evaluation methodology, so that in each phase of the planning problem both the data and the evaluation method can be critically judged.

Evaluation may relate to both sectoral planning issues (transportation network planning, facilities planning, e.g.) and integrated planning (comprehensive regional or urban planning, e.g.). In all cases, there should be a close agreement between information and evaluation.

INFORMATION SYSTEMS AND PLANNING

Information systems have a much broader scope than just a set of data. **Data** are only numerical representations of attributes of people, organizations, objects or events. **Information** may be defined as collection of organized data (for instance, by means of statistical techniques, modeling or transformation) so as to provide structured and systematic insight regarding a phenomenon. In this regard, an **information system** means any kind of systematic and coherent analytic or decision support system for planners and policy makers (cf. Rittel, 1982). Such a system serves to contribute to solving, organizing or rationalizing complex choice and decision problems.

Several **aims** of an information system may be mentioned (see Nijkamp and Rietveld, 1983):
- an aid to integrated or multidisciplinary analysis
- a contribution to operational and empirically-oriented scientific research
- an increase of the effectiveness of policy making and planning
- a contribution to building, testing and using practical models
- a rationalization of conflicting interests between groups or decision agencies.

It is evident that the design and maintenance of an information system is a far from easy task, as there are many costs involved, depending on the accuracy, adaptability and availability of the necessary information. On the other hand, the relevance of information systems have also to be judged on the basis of their benefits to improved decision-making (in terms of risk avoidance, higher effectiveness, multiple purpose use, or higher efficiency). Inadequate information may render economic planning models ineffective, may lead to a misinterpretation of socio-economic processes, may cause inconsistent or incoherent decisions, and may hamper the necessary communication between policy makers and experts.

In general, an information system aims at increasing our knowledge regarding a complex phenomenon. If new information leads to a decrease in certainty regarding the expected outcome of a decision, it is called "mis-information" (Rittel, 1982). In general, however, information brings more order to an otherwise less organized complex system, so that a rise in the information content removes uncertainty and reduces the entropy of a system ("negentropy"; see Scheele, 1983).

Information means a treatment and transformation of data. Examples of such operations are: capturing, verifying, classifying, arranging, summarizing, calculating, forecasting, simulating, storing, retrieving and communicating. Clearly, data can be transformed toward various aggregation levels (groups or regions e.g.). An aggregate representation of information implies a certain loss, which can be measured by the **entropy** of the system at hand: entropy measures the extent to which micro variables are ignored when one knows only a macro variable (cf. Gokhale and Kullback, 1978).

At any level of aggregation, information systems may in principle be used for three purposes (or stages) of policy analysis:

* **description**: a structural representation of a complex system (for instance, by means of multidimensional profiles, statistical tools or models)
* **impact analysis**: an assessment of effects of policy measures (for instance, by means of simulation models, qualitative effectiveness analysis, etc.)
* **evaluation**: an assessment of the merits of alternative courses of action (for instance, by means of cost-benefit analysis, multiple criteria analysis, etc.).

If one adopts a procedural view of planning, in which decision-making is regarded as a process, information systems have to be flexible, so as to provide at any desired moment decision agencies with specific tailormade information. This procedural view of planning leads thus in essence to the design of **adaptive** information systems, which have gained much popularity in recent years. This tendency runs parallel to the recent design of interactive user-oriented multiple criteria decision models (see, for instance, Nijkamp, 1980, and Rietveld, 1981).

Clearly, the needs of (private or public) decision agencies for information systems depend very much on the nature of the choice or policy problem at hand. These needs are **inter alia** determined by: the frequency of the choice problem, the range

of the impacts, the number of spillover effects, the number and intensity of policy conflicts, the financial implications, the time horizon, the number of decision agencies, the degree of uncertainty regarding the outcomes, and the degree of reversibility of the choice or policy problem at hand.

Clearly, may tradeoffs have to be made in designing appropriate information systems, such as the aim of a maximum accuracy of input data, the aim of maximum quality and usefulness of information systems, and the aim of the best possible treatment of a complex choice problem. The conflicts between these aims are represented in Figure 1, where the three-dimensional plane reflects the possibility frontier for arriving at a given quality of a choice to be made, when this choice is hampered by three barriers (viz., quality of input, quality of throughput, and quality of policy evaluation).

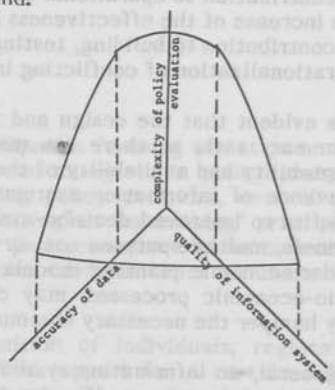


Figure 1. **Tradeoffs among three aims**

REGIONAL INFORMATION SYSTEMS

Especially during the seventies, local and regional governments became increasingly aware of increasing regional disparities and conflicts and of increasing urban decline. Policy measures to cope with such undesirable developments are necessarily to be based on reliable, consistent and up-to-date information, preferably at a disaggregate level. The design of a spatial information system, however, is a far from easy task, as a closer analysis of the geographic dimensions of a complex dynamic system requires the collection of a great deal of data, the use of an organized structure of the spatial system at hand, and adequate insight into horizontal and vertical patterns and processes in such a system. It is therefore no surprise that there is a great diversity in spatial information systems, such as: areal unit information, natural topological data, surface information, graphic symbol data, and label text information.

Especially in the seventies, several regional and urban information systems were developed (see Hermansen, 1971, and Willis, 1972), and the advances in computer technology have favoured systematic storage, processing and monitoring of large data sets in such systems. This development was also favoured by recent advances in carthographic techniques, such as color display and choropleth mapping (e.g. Tobler, 1979, and Steiner, 1980). In this regard, also geocoding has to be mentioned, as a geocoding system is a highly important tool for improving the organization of data and the display of information at a disaggregate scale (cf. Van Est and De Vroege, 1983).

A comprehensive survey of computer software for spatial information systems has been undertaken by the Commission on Geographical Data Sensing and Processing of the International Geographical Union (1980). In addition to a systematic presentation of various spatially-oriented information systems (such as full geographic information systems, data manipulation programs, and carthographic and graphic techniques), also an "ideal" geographic information system was

described. This system compromises six major subsystems, viz. management, data acquisition, data input and storage (control processes, encoding, filing, etc.), data retrieval and analysis (comparitive and statistical analysis, e.g.), information output, and information use (user-expert dialogue, man-machine interaction, e.g.). This approach clearly indicates that spatial information systems should not just act as data bank systems, but - more importantly - as tools for planning device.

An evident problem inherent in designing a spatially-oriented information system is the regionalization and hence the level of **spatial aggregation** of the variables taken into account (for instance, zones, urban districts, etc.). The problem of areal demarcation cannot be solved unambigously, as any regionalization can be based on various viewpoints, such as institutional-administrative principles, functional-economic principles or statistical principles. Clearly, if an information system contains a wide variety of variables, one has to apply also **clustering** and **classification** principles, so as to design an information system that reconciles the diverging options for various components (or variables) in an information system. In this respect, the recently developed **geocoding** techniques offer a great potential, as they are able to combine the aim of any appropriate spatial detail with the aim of flexibility in spatial aggregation.

In general, any regional or urban information system should serve multiple purposes so as to provide the necessary tools for integrated regional or urban planning. Therefore, the following criteria may be mentioned which may serve as a general framework of reference for judging an adaptive information system: accessibility, consistency, completeness, rapid availability, timeliness, policy relevance, pluriformity, comparability with other information, flexibility, measurability of information, integration, multi-purpose nature, and statistical validity. Furthermore, some more specific geographic aspects of spatially-oriented information systems have to be mentioned, viz. vertical and horizontal coherence, degree of spatial interaction, existence of spatial spillover effects, identification of specific bottlenecks in regional or urban development, inclusion of multiregional decision making aspects, and meaningful statistical standardization of data.

Several earlier examples of regional information systems can be found among others in Hägerstrand and Kuklinski (1971), Kuklinski (1974), Perrin (1975), Benjamin (1976), Guesnier (1978). More recent applications can be found among others in Elfick (1979), Peters (1981), Kitamura (1982), Petzold and Heineke (1982), Van Est and De Vroege (1983) and Scheele (1983). There appears to be an increasing tendency to link spatial information systems with automated computer-assisted systems (for instance, plotter-drawn maps, spatial diagnostic checking methods, and regional and urban modeling). Clearly, spatially disaggregate information can be used at various spatial aggregation levels, pending on the specific analytical or policy interest (cf. Issaev et al., 1982, and Nijkamp and Rietveld, 1983). If one makes a distinction between the data input, a systems model and the information output of an information system, the use of information in a spatially disaggregate system can be represented by the following figure (Figure 2).

In the recent past, many specific tools for information systems have been designed to cope with uncertainty in regional and urban information systems. Some major tools in this respect are:

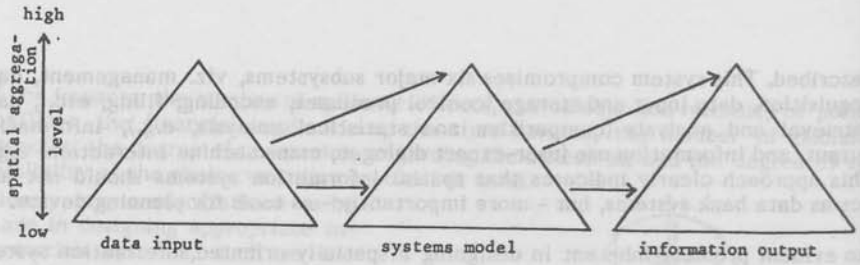


Figure 2. **Structure of a spatially-oriented multi-level information system**

* **Scenario analysis**: This approach serves to cope with long-term uncertainty regarding the structure of complex and dynamic spatial planning problems.

* **Early warning systems**: This method aims at designing the critical moments a certain policy should come into action.

* **Qualitative evaluation analysis**: This method tries to identify the most reasonable course of action for a choice problem with conflicting interests and qualitative information.

* **Decision support systems**: This approach serves to improve the quality of decision making by means of man-machine interactions (or computer consulting systems) regarding qualitative and less structured decision problems.

* **Monitoring**: This decision aid technique attempts to store and retrieve data on spatial processes with a specific view of planning processes.

Despite many advances in spatial information systems, in almost all countries many elements are still lacking in the design, contents and use of such systems. Examples of such missing information are: stocks and flows of wealth, capacity constraints, disaggregate spatial interaction flows, regional and urban innovation efforts, institutional patterns, and long-run regional dynamics. Clearly, there is much variety in quality and quantity of spatially-oriented information systems among various countries. This has been elaborated in Nijkamp and Rietveld (1983).

A PERSPECTIVE

A recently performed international comparitive analysis of spatially-oriented information systems illustrates that, despite the available technology (computers, data processing techniques, etc.), the current state-of-the-art in regional information systems is not very advanced (cf. Nijkamp and Rietveld, 1983). It is particularly suprisong that many **multiregional and multinational companies** have managed to build ip appropriate information systems regarding all aspects which are considered to be relevant for their decision making, and in this regard, regional information analysts might learn many lessons from large business companies.

The design of adequate spatial information systems requires satisfactory insight into the structure, evolution and mechnisms of regional development. In this regard, **key factor analysis** may be an important tool, as this approach aims at identifying the driving forces (or key factors) of a complex dynamic system. Useful tools in this respect are: causality analysis, graph theory, qualitative calculus and bifurcation theory (see also Brouwer and Nijkamp, 1983).

Another problem concerns the matching of information systems to administrative units. In a multilevel spatial structure, it is in general desirable to build information systems in a **bottom-up fashion** so as to let them fit flexible into any desired level of spatial planning and policy making (cf. McDowell and Mindlin, 1971). Consequently, much attention has to be devoted to segment-based spatial reference systems or refined grid or zoning systems. The geocoding system is a good example of this.

The integration of diverse components in a spatial information system is another major problem. In this respect, it is meaningful to employ a **satellite** principle. This implies that a information system is built up by first including the key component which is relevant for the policy process at hand. Then new components are added in as far as they are related to the kernel of the information system or in as far as they are necessary for a multipurpose information system (see Figure 3).

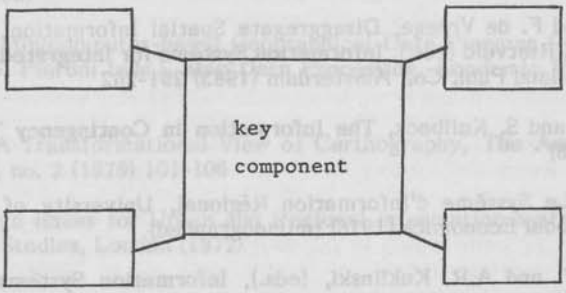


Figure 3. **A satellite structure of information systems**

Evidently, in many cases the accessibility and use of information systems is hampered by **confidentiality** requirements. Confidentiality of business data however, is a relative concept, as in general competitors know each other's market position, technology, and investment plans quite well. In many countries, there is a tendency to overestimate the confidentiality requirement. In this regard, one may adopt the viewpoint that information for public planning is a public good, and that the quality of planning (for both public and private aims) may be enhanced by putting less emphasis on confidentiality of information for planning.

Finally, there is a need for more **user-friendly information systems** (cf. Mayer and Greenwood, 1980, and Sol, 1983). In various cases, information systems contain much information for public planning, but several policy making agencies have great difficulties in understanding the structure and mechanism of large information systems (including modeling). In this respect, user-friendly automated computer systems (e.g., interactive computer graphics, desk-top computer terminals) may pave the road toward bridging the gap between information experts and responsible policy agencies.

REFERENCES

Benjamin, B., Statistics and Research in Urban Administration and Development. International Association for Regional and Urban Statistics, The Hague (1976)

Brouwer, F. and P. Nijkamp, Qualitative Structure and Stability Analysis. In: P. Nijkamp, H. Leitner and N. Wrigley (eds.), **Measuring the Unmeasurable**, Martinus Nijhoff,Dordrecht(1985)

Burch, J.G., F.R. Strater and G. Grudnitski, **Information Systems: Theory and Practice**, Wiley, New York (1979)

Commission on Geographical Data Sensing and Processing, International Geographical Union, Ottawa (1980)

Debons, A. and A.S. Larson (eds.), **Information Science in Action**, Martinus Nijhoff, The Hague (1983)

Elfick, M. (ed.), **URPIS Seven**, Melbourne: The Australian Urban and Regional Information Systems Association (1979)

Est, J. van, and F. de Vroege, Disaggregate Spatial Information Systems. In: P. Nijkamp and P. Rietveld (eds.), **Information Systems for Integrated Regional Planning**, North-Holland Publ. Co., Amsterdam (1983) 251-262

Gokhale, D.V. and S. Kullback, **The Information in Contingency Tables**, Dekker, New York (1978)

Guesnier, B., Le Système d'Information Régional, University of Poitiers: Institute of Regional Economics (1978) (mimeographed)

Hägerstrand, T. and A.R. Kuklinski, (eds.), **Information Systems for Regional Development**, University of Lund, Lund, Sweden (1971) 1-37

Hermansen, T., Information Systems for Regional Development Planning. In: T. Hägerstrand and A.R. Kuklinski, eds.), **Information Systems for Regional** Development, University of Lund, Lund, Sweden (1971) 1-37

Issaev, B., P. Nijkamp, P. Rietveld, and F. Snickars (eds.), **Multiregional Economic Modeling: Practice and Prospects**, North-Holland Publ. Co., Amsterdam (1982)

Kitamura, T., An Agricultural Model in the Regional Framework, Interim Report, International Institute for Applied Systems Analysis, Laxenburg (1982)

Kuklinski, A.R. (ed.), **Regional Information and Regional Planning**, Mouton, The Hague (1974)

Mayer, R.R. and E. Greenwood, **The Design of Social Policy Research**, Prentice Hall, Englewood Cliffs (1980)
McDowell, B.D. and A. Mindlin, Obtaining Metropolitan Planning Data from Local Governments. **Journal of the American Institute of Planners**, March (1971) 111-115

Nijkamp, P., **Environmental Policy Analysis**, Wiley, Chichester/New York (1980)

Nijkamp, P. and P. Rietveld (eds.), **Information Systems for Integrated Regional Planning**, North Holland Publ. Co., Amsterdam (1983)

Perrin, J.C., **Le Développement Régional**, Presses Universitaires de France, Paris (1975)

Peters, A., Die Verstätterung in den Niederlanden nach 1950 und seine Raumordnungspolitischen Implikationen, Ph.D. Thesis, Technical University, West-Berlin (1981)

Petzold, E. and H.J. Heineke, Geographic Information Systems, Report, Institute of Geography, Westfalian Wilhelmus University, Münster (1982)

Rietveld, P., **Multiple Objective Decision Methods and Regional Planning**, North-Holland Publ. Co., Amsterdam (1981)

Rittel, H., Structure and Usefulness of Planning Information Systems. In: P. Laconte, J. Gibson and A. Rapoport (eds.), **Human and Energy Factors in Urban Planning**, Martinus Nijhoff, The Hague (1982), 53-64

Scheele, R.J., Making Plans Work, Ph.D. Thesis, State University, Utrecht (1983)

Sol, T.W. (ed.), **Processes and Tools for Decision Support**, North-Holland Publ. Co., Amsterdam (1983)

Steiner, D., A Minicomputer-based Geographical Data Planning System. In: C.H. Freeman and G. Pieroni (eds.), **Map Data Processing**, Academic Press, New York (1980) 1-25

Tobler, W.R., A Transformational View of Carthography, **The American Carthographer**, Vol. 6, no. 2 (1979) 101-106

Willis, J., Design Issues for Urban and Regional Information Systems, Centre for Environmental Studies, London (1972)

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EVALUATION BY GRAPHICS

R. Daru
Faculty of Architecture, Building and Planning
Eindhoven University of Technology

INTRODUCTION

Analysis and evaluation techniques offer the means for ordering, structuring and judging the decision alternatives identified within the planning process. But the techniques presently being applied or developed could be made more fruitful and efficient if more attention were given to their graphical presentation. The graphic prosessing of information provides more insight into the problem being studied, making evident that which is sometimes obscured when presented in alpha-numeric tables or algebraic representations. What is there, but hidden, is revealed.

But there is more. The collected data can be made meaningful by graphic processing and presentation. In their graphic form, data can suggest new patterns which lead to the answers to questions one wasn't even aware of to begin with. However, this implies 'graphic thinking', or 'graphicacy', as Balchin and Coleman (1) would put it. Graphicacy means more than beautifying given information by graphic means. Neither is it the duplication of succesfull examples, or the application of recipes. Jacques Bertin, whose magnum opus 'Sémiologie graphique' of 1967 has just been translated into English (2), is one of the few researchers who have dedicated himself to the development and implementation of the principles of graphicacy. Bertin is the first author who have created a consistent system based upon the still unsurpassed ability of the human eye to grasp differences and patterns in graphically offered data structures.

Bertin's background is geographical an cartographical. His approach was born out of the shortcomings of graphical means within his discipline, in particular those stemming from thematic maps meant to classify and cluster data in a spatial way. The usual cartographic techniques fell short of the goals set by the geographer. Bertin was then confronted with the choice between using multivariate analysis techniques, or finding a new means of cartographical implementation. He chose to make use of his insight and knowledge of cartography and data-analysis techniques to make an original path. His results do not have to be confined to geographical applications. Besides data-exploration for all kind of research problems, they can help planners and designers to support their decision-making.

The theory, method, and practice of the Bertin approach will be sketched here, primarily demonstrated by means of examples (3).

GRAPHICS PERCEPTION AND EVALUATION

Graphics (used here to mean theory as well as results) is based on the assumption that every drawn or printed image is composed of marks, and that its perceptual properties must correspond with the measurement levels represented in the data to be presented. A consequent application of this principle ensures that the perception of the relevant information will occur with sufficient immediacy,

immediacy being the one criterion for adequate graphical presentation of data. This principle, and this criterion, have lead to various specific applications of graphical techniques. Although the Bertin techniques reflect their origins, the principle and criterion apply to all graphical images, not only those used in cartography. This is why a number of evaluation techniques for policy analysis will be applied here, according to the above named principle and criterion, using graphical means to broaden and refine them.

According to Bertin the 'semiology of graphics' - or science of graphical sign systems - teaches that data properties cannot be presented by whatever kind of graphical mark is available, but that data property and graphical property must harmonize. For example, when **quantities** have to be transcribed, this must be done by using the **size** (or rather magnitude) of the elementary marks. They must moreover, conform to the perceptual law of constant proportionality. This means that the quantitive progression of the marks must be bases on their area (4).

Apart from by size, **ordinal arrays** can be made evident by the use of **value** (from black to white via grey or vice versa) and/or **grain** (from coarse to fine or vice versa). Grain and value can be realized graphically by the afore mentioned means, plus by variation of **colour** and **orientation** (for example the inclination of a bar). Least differentiated perceptually is the **form** of graphical signs. All other things being equal, form variation does not help much in seeking patterns (see Figure 1.2.5).

THE ASSESSMENT MATRIX

To illustrate the use of the six basic graphic variables, we shall look at the example of an assessment matrix (presented in its original form in Figure 1.1). (5) To be able to transform the numerical data directly into graphical signs in a logical manner, we have to change the ranking (1st, 2nd, 3rd) (Fig. 1.1) into a score, with the best of the three alternatives getting the highest score (3 for the best, 2 for the second best, 1 for the worst). We then proceed to transform best into (graphically) heaviest. We can do this by using value (Fig. 1.2.1) or grain (Figure 1.2.2). With value or grain, the ranking of alternatives can be perceived spontaneously, immediately. If we look at Figures (1.2.3, 1.2.4 and 1.2.5), this immediacy is absent. We have to concentrate, reflect, and consult a legend if we are to grasp the intended order.

1.1 Assessment matrix with ranking integers

	Alternatives		
Criterion	A1	A2	A3
C1	1	2	3
C2	2	1	3
C3	3	1	2
C4	3	2	1
C5	2	3	1

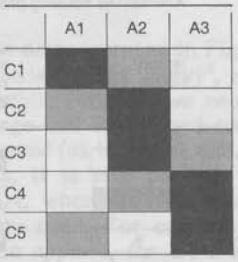
For the sake of the example, the original numbers have been replaced by quantities (Figure 1.3) which can be translated back into three classes. In Figure 1.3.1 we make use of the matrix cells to express the graphic variable of size, whereby the highest quantity corresponds to the complete area of one cell, and the other quantities are in proportion.

Combinations of graphic variables must be used with caution. If we return to the data as given in Figure 1.2 and use shape, orientation, and value combined (Figure 1.4.1), we weaken the effect obtained, although still using graphic levels adequately in correspondence with measurement levels (the ranking is still present).

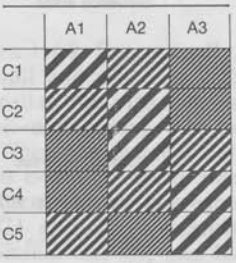
1.2 Reversal of 1.1

	A1	A2	A3
C1	3	2	1
C2	2	3	1
C3	1	3	2
C4	1	2	3
C5	2	1	3

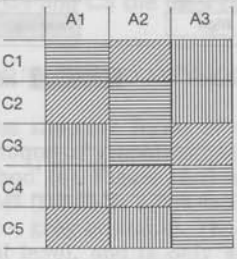
1.2.1 Value



1.2.2 Grain



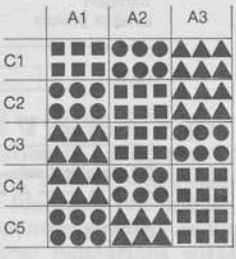
1.2.3 Orientation



1.2.4 Colour

	A1	A2	A3
C1	red	blue	yellow
C2	blue	red	yellow
C3	yellow	red	blue
C4	yellow	blue	red
C5	blue	yellow	red

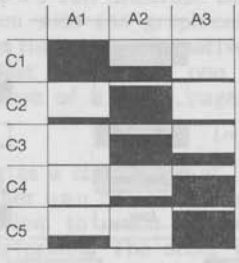
1.2.5 Shape



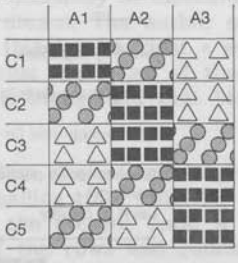
1.3 Quantification of 1.2

	A1	A2	A3
C1	103	38	7
C2	17	95	11
C3	2	75	29
C4	5	24	78
C5	31	3	90

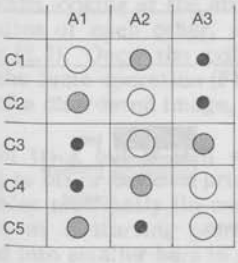
1.3.1 Size



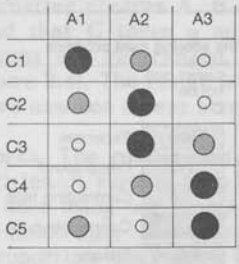
1.4.1 Comb. of mark patterns



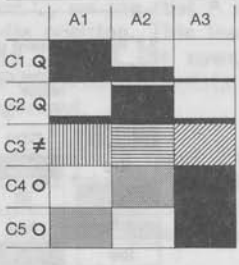
1.4.2 Neg. combination



1.4.3 Pos. combination



1.4.4 Comb. of meas. levels



2.1 Original visualisation of the scoring card technique

Influence areas	Alternatives housing provision		
	rented house A	owner occupied house B	owner occupied house C
Monthly expenses	600	1200	900
Living space (in m²)	100	150	110
Commuting time (in min.)	5	70	30
Connection to public transport	medium	bad	good
Level of privacy	low	high	medium

Rank: best medium worst

* All numbers are fictitious

2.2 Graphical value reversal

Influence areas	Alternatives housing provision		
	rented house A	owner occupied house B	owner occupied house C
Monthly expenses	600	1200	900
Living space (in m²)	100	150	110
Commuting time (in min.)	5	70	30
Connection to public transport	medium	bad	good
Level of privacy	low	high	medium

Rank: best medium worst

2.3 Rearrangement of the scoring card columns

rented house A	owner occ. house B	owner occ. house C	
600	1200	900	Expenses
100	150	110	Living space
5	70	30	Commuting time
medium	bad	good	Connect.to publ. transp.
low	high	medium	Privacy

best
medium
worst

owner occ. house B	owner occ. house C	rented house A	
1200	900	600	Expenses
150	110	100	Living space
70	30	5	Commuting time
bad	good	medium	Connect.to publ. transp.
high	medium	low	Privacy

The worst possible combination is demonstrated in Figure 1.4.2. Perceptually, both black and large elements are experienced as 'heavy', small and white as 'light'. If we combine large (=heavy) with white (=light), we neutralise the visual impact. If on the contrary, we combine large and black (and small and white), the ranking effect is maintained, if not enhanced (as in Fig. 1.4.3). When the original data are of differing measurement levels, it is still possible to combine them into one image. This is done in Figure 1.4.4, where various graphical means, corresponding with each measurement level, are used. For criteria C1 and C2, a quantitative graphical translation (Q) has been applied, for C3 a nominal one (≠), and for C4 and C5 an ordinal one (O). For C1 and C2, the data from Figure (1.3) have been used, for C4 and C5 the data come from Figure 1.2.

THE SCORING CARD METHOD

One of the simplest forms of using graphical variables is shown in Figure (2.1). This technique of graphical data prosessing and presentation is called the 'scoring card method' (6). Figure (2.1) is the original presentation. It uses shape and value differences (shape as alpha-numeric signs, and value as identical rectangles filled with white, grey or black). The quantitative level has been graphically reduced to an ordinal level, and is only stil visible in the numerical indications within the rectangles. The author seems conscious of the lack of spontaneous ordering on the basis of the reading of numbers, and has used value as a perceptual support. With this in itself simple presentation, we can illustrate the necessity of matching the aim of the graphical presentation with the graphical means. The author of the original presentation accentuates the worst alternative (black = heaviest = worst). If one wishes to pick out the best alternative, one must make use of the best perceptual value, which in the case of a white page is black, and thus black = heaviest = best (see Figure 2.2).

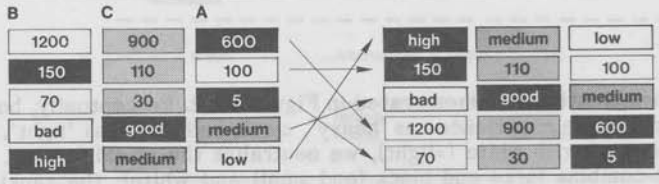
If one takes the ease of reading as a starting-point, then one wishes to present him/her with the results of his/her own choice. The reading time must be minimalised without loss of information. to reach this in the present case, a small number of operations can be performed. The order of the rows and columns in Figures 2.1 and 2.2 is arbitrary. Consequently, the information can be rearranged at will. When looking at the alternative choices A, B and C, we see that A and B are opposites of each other, and that C takes a middle position. This leads to operation (2.3). Once the columns are rearranged, we look at the rows, and perform the same operation (Figure 2.4). The result is a diagonalised matrix (2.5). In place of a disordered image, we now see a neat arrangement.

It's an old trick but it still works. The reader can now make a quick choice, according to his or her own priorities. A further step can be taken by quantifying the variables metrically (in our case using bar length). the maximal data present are taken as a starting point (maximum = longest bar). The other data are translated into smaller bars in proportion (see Figure 2.6).

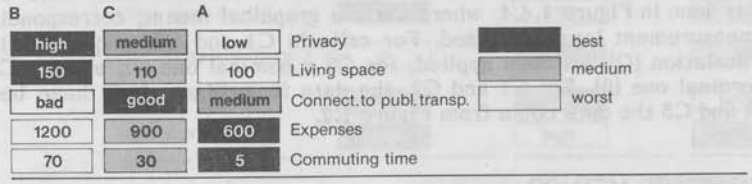
If the variables vary in relative importance, the bars can be redrawn according to the given weights (Figure 2.7). We can then proceed by taking each of the alternatives in turn as a standard. In (2.8) the rows of A are rearranged from the smallest to the largest bar, and the rest follows. The same is done in (2.9) using B as a standard. We can see that C keeps in both cases a profile more similar to A than to B (which is why both are shaded to enhance this fact).

If we want to emphasize deviations from a middle position, we can make use of a zero-line. The middle position for each of the variables is then represented by the

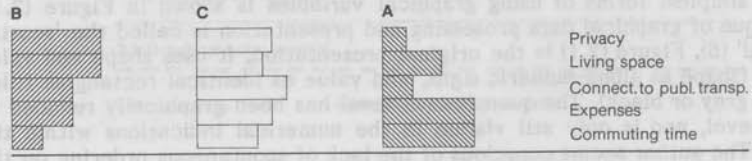
2.4 Rearrangement of the scoring card rows



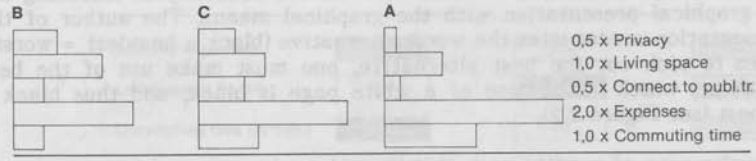
2.5 Result of the rearrangement (see 2.3 and 2.4)



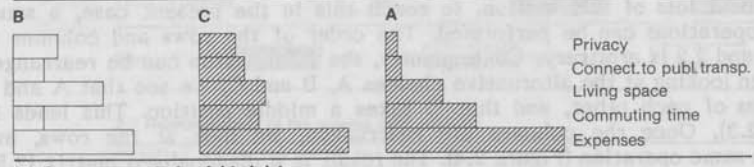
2.6 Quantitative expression of scoring card data



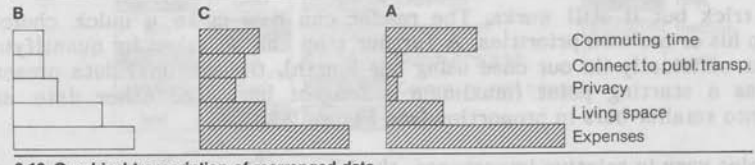
2.7 Data of 2.6 in weighted form



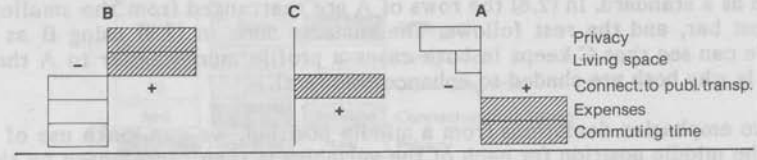
2.8 Rearrangement of weighted data on the basis of A



2.9 Rearrangement of weighted data on the basis of B

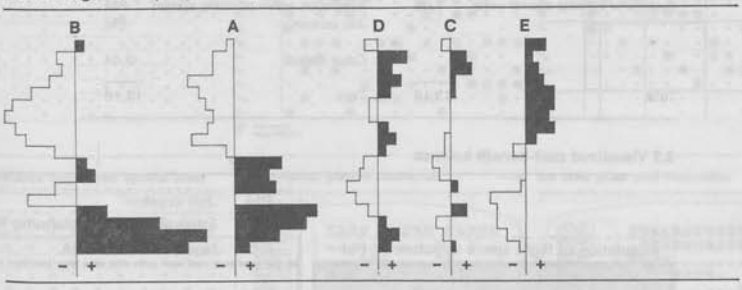


2.10 Graphical transcription of rearranged data of (2.5) around a zero line



zero-line, while positive and negative (above or below middle) data, as shown in Figure (2.5), are placed on each side of the line (see 2.10). Since C represents the middle position in 4 out of 5 cases, it has only one bar (Note that this is not a statistical representation in any strict sense). In fact, the data in (2.5) do not necessiate such an arrangement. In the case of a greater number of alternatives and variables however, the necessity of such a representation makes itself felt (as in 2.11). In such a case, diagonalisation is even more adequate as a means of clarification.

2.11 Enlarged example (ref. 2.10)



COST-BENEFIT ANALYSIS

Even in the case of cost-benefit analysis (7), where balance sheets seem clear enough, graphical representation can bring a perceptible emphasis which is not present in the words and numbers presentation, as is made evident in (3.2) as compared to (3.1). The large sums can be fully appreciated in comparison to the small ones, and the surplus of the balance is emphasized, while at the same time its relative importance can be felt with sufficient immediacy. The quantities surrounded by broken lines represent the unknowns. This single graphical representation expresses the global situation, while retaining the detailed information of the balance sheet. The detailed information cannot be included with the same ease when handling a larger number of alternatives (see 3.3). In this case, the sums can still be translated with some rigour in a proportional dot matrix representation. As in the previous cases, a first version can be submitted to rearrangement, and the diagonalised result can lead to the perception of clusters of alternatives.

THE PERMUTATION METHOD

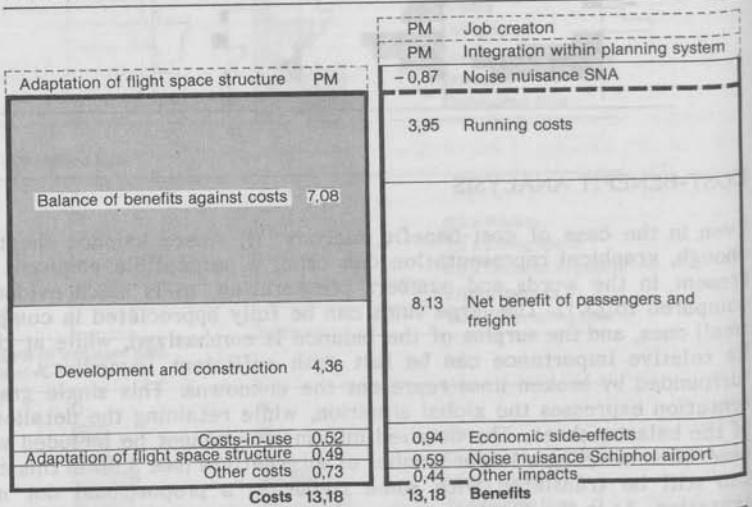
In the previous examples, we have already made use of permutation, but in the context of other techniques. In the case of the Jacquet-Lagrèze method (8), permutation is not only central to the graphical presentation as outlined here, but also to the method itself. Figure 4.1 shows us the first and second priorities of the members of a committee. The predominance matrix in Figure 4.2 is the result of calculations taking into account all the choices made among the alternatives, by all the members. Visually speaking, all the resulting numbers (which in themselves are the result of a number of permuting operations) are more or less equivalent. The only immediately perceptible items are the ties. The ordering from top to bottom and from left to right represents the resulting order from best to worst.

3.1 Summarised cost-benefit balance for a second national airport (SNA)

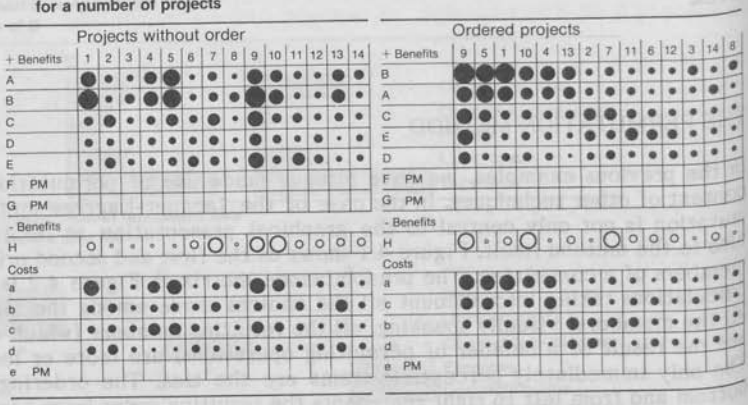
Amounts in 10^9 guilders, price level of 1973
Period 1985-2015
Discounted value 10% per annum after 1985

Costs		Benefits	
Development and construction costs	4,36	Running costs	3,95
Costs-in-use	0,52	Net benefits of passengers and freight	8,13
Adaptation of flight space structure	0,49 +PM	Economic side-effects	0,94
Other costs (including transport infrastructure)	0,73		
		Noise nuisance SNA	–0,87
		Noise nuisance Schiphol airport	0,59
Balance of benefits against costs	7,08	Integration within planning system	PM
		Job creaton	PM
		Other impact	0,44
Total	13,18	Total	13,18

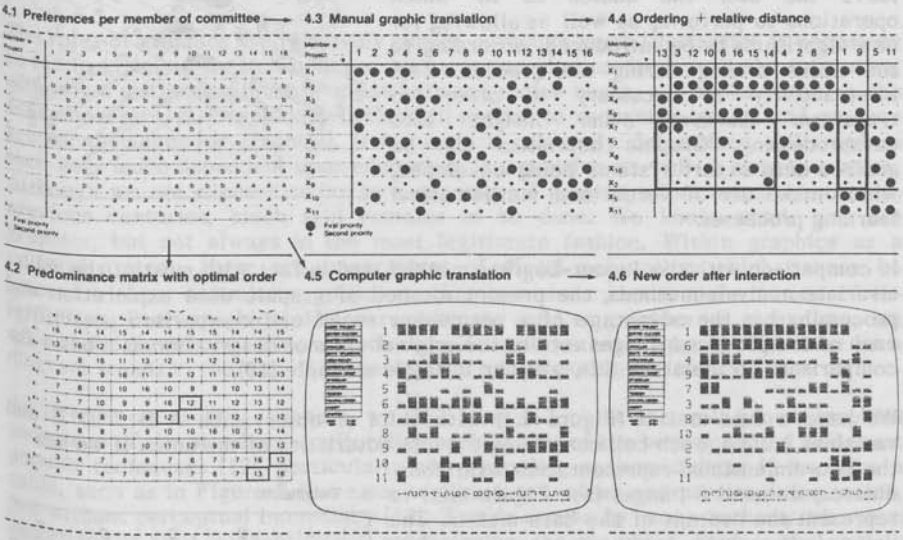
3.2 Visualized cost-benefit balance



3.3 Visualized cost-benefit balance for a number of projects



A direct graphical processing of the data-matrix in Figure 4.1 by manual means leads to Figure 4.3. A reshuffling on the basis of visual impressions results in Figure 4.4. The rank-order is identical with the Jaquet-Lagrèze rank-order, but with the added benefit that more information can be induced from the graphical matrix than from the numerical matrix. The distance between projects and between judges can be perceived.

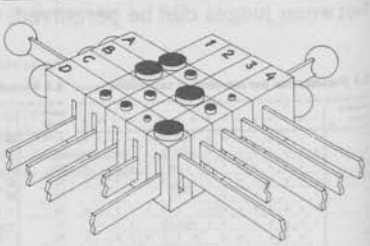


We can see for example, that the projects X1 and X6, which are nearest neighbours, are almost equivalent, whereas the nearest neighbours X4 and X9 differ to a much larger extent, with symmetrical priorities from the judges. The group of projects X7, X3, X8, X10 and X2 all score rather poorly, and get very scattered preferences, whereas X5 and X11 almost unanimously get poor marks. By keeping the data-matrix as such (instead of resorting to a predominance matrix), we have retained information about the judges. We can, for example, see that judges 1, 9, 5 and 11 are idiosyncratic, but each in a different way; 5 and 11 differ most from the rest, and seem to rank projects in a reverse manner. Judges 13, 3, 12, 10, 6, 16, 15 and 14 are, as a group, most affirmative for the upper four projects, whereas the preferences of judges 4,2,8 and 7 are less marked. The present example has a modest span, and the manipulation of rows and columns can be performed with paper, scissors and glue. when the matrix is around 100 by 100, a so-called manual permutator (9) can be used (see Figure 5), made of small plastic blocks with slots in which needles can be inserted. On the visible top of the blocks, a black dot protrudes. There are ten sizes of dots, representing proportionally marks from 1 to 10. The bottom side carries complementary dots (or no dot in the case of upper dot 10). There is also a special sign for missing or inapplicable data. By extracting the pins carrying the columns, the rows can be resorted, and vice-versa. Each step in the sorting process can be recorded by simply copying the obtained matrix on a standard copying machine. If, in the case of binary (yes=1=black; no=0=white) answers, the negative image can provide more insight by simplifying the resulting pattern of dots, the column or row can be turned over. When using the complementary dots, the underlying concept must also be reversed. This can be done by replacing the presence of a property by its

absence as the quality being examined.

Computer versions (Figures 4.5 and 4.6) follow the same principles, but quicken the sorting and reshuffling. A good programme ought to make the sorting operations, previously done by manipulating pins, automatic. It should also leave the user the choice as to which operations to perform, as well as allowing for new algorithms to be introduced, according to the needs arising during the process. This inter-activity is necessary in order to transform discovery and insight into dataprocessing. Graphic heuristics can be grafted onto expert systems being developed, which make use of verbal and mathematical learning processes.

5 Manual permutator (Detail 1:1)



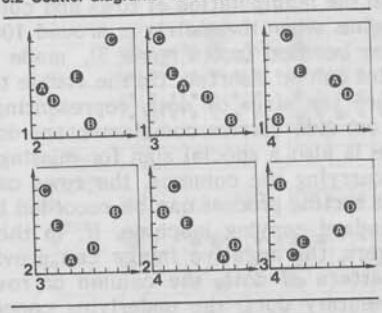
In comparison with the Jaquet-Lagrèze method, and in fact with most of the multivariate analysis methods, the present method of graphic data exploration and processing has the advantage of a permanent recall and comparison possibility: each of the produced images retains the original units of observation as a basis for comparison. To illustrate this, we take a simple example (10).

We have a data-matrix (Figure 6.1) with units of observation A B C D E and variables 1 2 3 4, each cell containing the characteristics of the units. If we use the two-dimensional representation wtih coordinates, we need 6 images (see Figure 6.2.) to represent the content of the data-matrix. The units of observation are then anonymous points in scatter diagrams. If we want to compare say C with A, we would have to perform some intricate operations. The multiple bar-chart representation in Figure 6.3 relieves us from this task. Each unit, and each variable, retains its identity, while the graphic image carries the essence of the information contained in the data-matrix. Moreover, rows (variables) and columns (units) are rearranged so as to produce patterns.

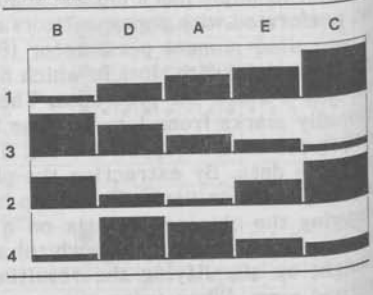
6.1 Datamatrix

	A	B	C	D	E
1	4	1	8	3	5
2	1	5	7	2	4
3	3	7	1	5	2
4	6	1	2	7	3

6.2 Scatter diagrams for 4 variables



6.3 Multiple bar chart



The first presentation (Figure 6.2) is geared towards the discovery of correlations between variables, the units as such are of no importance. The second representation (Figure 6.3.) has a double aim: to look at the relationships between variables, and to compare the units of observation. This second aim corresponds with the requirements of decision-making.

GRAPHICACY

The rules of graphics are relatively simple, although much illustration is necessary to aid in grasping them. Complete books are needed to display all the subtleties, whereby a distinction must be made between recipe books, and books based on the fundamentals (11). Graphicacy is the skill needed to recognise, grasp, and produce graphical images (or graphics). It is partly a culturally determined skill, and as such very much dependent upon education, in a positive as well as negative sense. Although we are exposed today to a quantity of graphic information unknown to previous centuries, much still remains to be done. We know, read, and use graphics, but not always in the most legitimate fashion. Within graphics as a cultural system, there are imperatives of visual perception which cannot be ignored without committing mistakes, that is without loss or deformation of information. Even graphical designers, who are used to graphics as a medium, sin against the 'laws' of graphics. Much too often, the esthetic appeal of a chart or a diagram stands in the way of its perceptual and cognitive efficiency.

As an example of what is meant, we shall have to look at the well known graphics technique of the pie chart - culturally speaking one of the best accepted and most popular techniques (12), particularly when handling economic data. If we take a table, such as in Figure 7.1, we see well ordered, clear and precise information, bur without perceptual immediacy (13). This is sensed by many authors, who then resort to the most 'evident' visual presentation, the pie chart. Such is the case in Figure (7.2), taken from a study on the Dutch building industry. If there is only one series of data - say the production of the building industry in 1975 - a pie chart would be an improvement upon the alpha-numeric information. The proportions could immediately be perceived, provided that the graphical translation is right in terms of pie proportions. This is not the case here! (see Figure 7.3). The graphical sin is committed by using two pie charts: the reader's eye must constantly leap from one pie to the other to make comparisons. Why not keep the original table as it was? Sellers of computer graphics packages still make the same mistakes, both in three dimensions (Figure 7.4), and in colour for good measure, and they even make you pay for it...

In Figure 7.5, another graphical presentation is shown. Of course it does not have the familiar look of the pie chart. But with a little training in graphicacy, it can immediately give answers to a number of questions; like the relative importance of the various sectors of the building industry in both years, which sector diminishes, stagnates, or grows, and how much. A reader can select the answer most appropriate to his/her position in the building process, and act accordingly.

GRAPHICAL MANIPULATION OR DIRECTED PERCEPTION ?

Graphics is a system of signs, and as with all other systems of signs, it can be made to lie. This is one more reason for training in graphicacy. Only with sufficient knowledge of the medium is it possible to see through willfull or unintended manipulation. An example is the reaction of one of the persons responsible for the report containing Figure (7.2), claiming that the small Figures were so uncertain that is was better not to draw attention to them, and that as such,

Figure (7.2) was preferably to Figure (7.5).

As it is well known, rates of change have to be represented by using a logarithmic scale (see Figure 8.2). When comparing the rates of growth over a number of years, only a semi-logarithmic chart will indicate exactly the differences and similarities between two or more curves. Consider authors who present the reader with an arithmic chart (like Figure 8.1) meant to do the same. Is this lack of graphicacy or lack of ethics ?

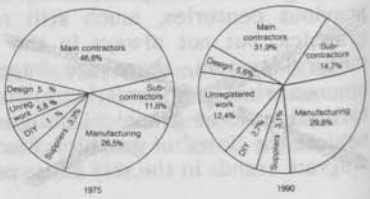
Classical 'How to lie with statistics' tricks still find a market among graphical mediators; tricks like 'the dissapearing zero-line' (Figure 9.1) or 'the treacherous broken scale' (Figure 9.2), or 'the compressed time scale', are always good for a bit of dramatization. What is really bad about these tricks, is that they can lead to wrong interpretations and inadequate decisions. It is true that the choice of scale units is arbitrary, but it should not lead to the disappearance of unevennesses in a curve, or to a lack of readability. There are, in fact, perceptual habits which ought to determine the choice of graphical means, related to the criteria of visibility of difference, verifiability, and plausibility.

Within the margins delineated by these criteria, one sometimes has to emphasize a number of graphical elements in order to demonstrate what one is talking about. One example of such a case, is when marginal differences are essential to one's hypothesis or aims. Take the problem of conservation of rare plant species: a common dot scale as applied for Figure 10.1 - which should be read as three maps of the same area - makes the distribution of rare plants in Figure 10.1.3 scarcely visible as compared to the more common species in the same area in Figures 10.1.1 and 10.1.2. When using a separate dot scale per map, as in Figure 10.2, the concentration of rare plants is visibly enchanced, and the determination of areas to be protected is made easier and more precise. In such a case, where the aim of the graphical representation is perfectly clear, such an emphasis is justified.

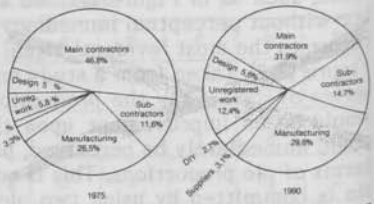
7.1 Market value of building production in added value and other costs

	1975	1990
Main contractors	46,8 %	31,9 %
Manufacturing	26,5 %	29,6 %
Sub-contractors	11,6 %	14,7 %
Unregistered work	5,8 %	12,4 %
Design	5 %	5,6 %
Suppliers	3,3 %	3,1 %
Do-It-Yourself (DIY)	1 %	2,7 %

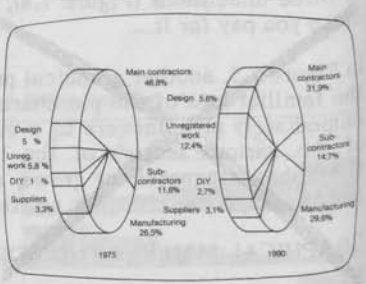
7.2 Inadequate original visualisation



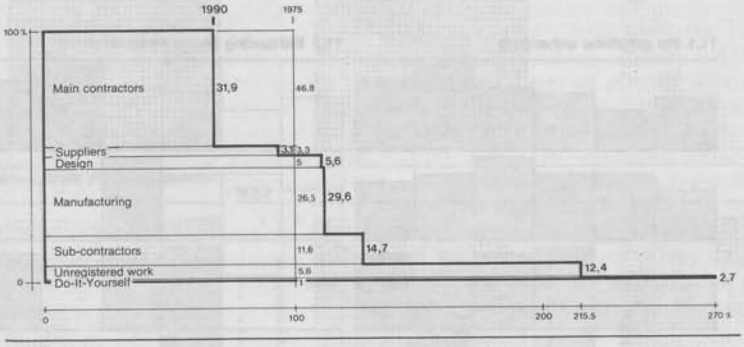
7.3 Corrected original visualisation



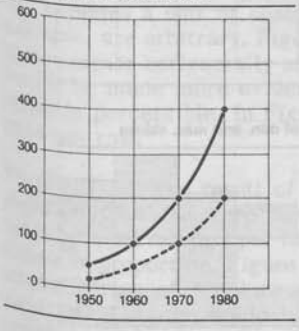
7.4 Computer graphic visualisation



7.5 Improved visualisation

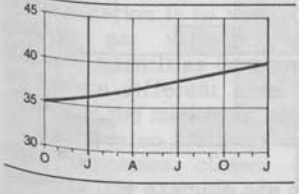


8.1 Line chart with equal spacing

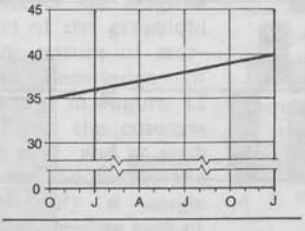


8.2 Semi - logarithmic scale division

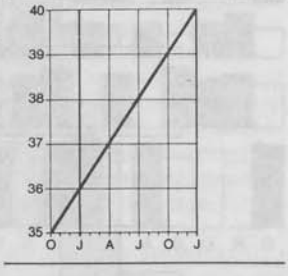
9.1 Line chart without zero line



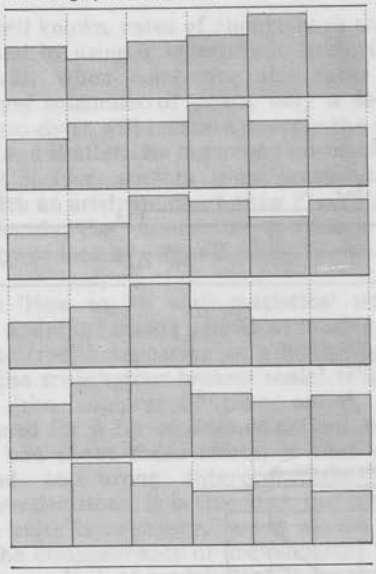
9.2 Line chart with break in vertical scale



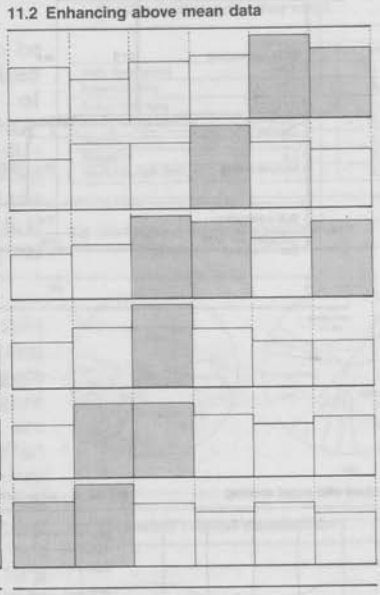
9.3 Line chart with changing scale division



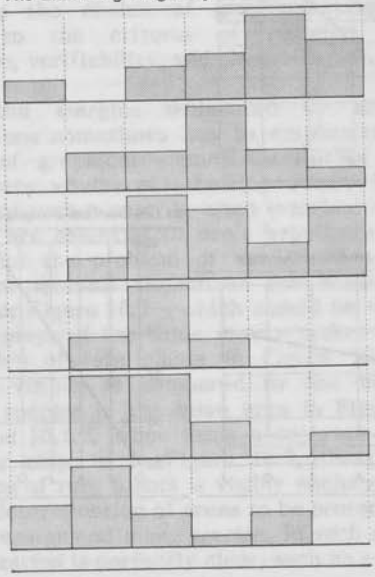
11.1 No graphical enhancing



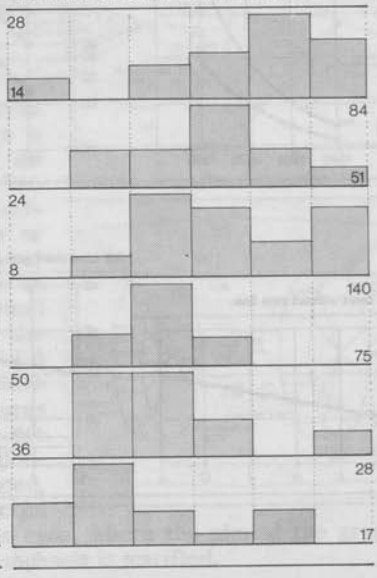
11.2 Enhancing above mean data



11.3 Enhancing marginality

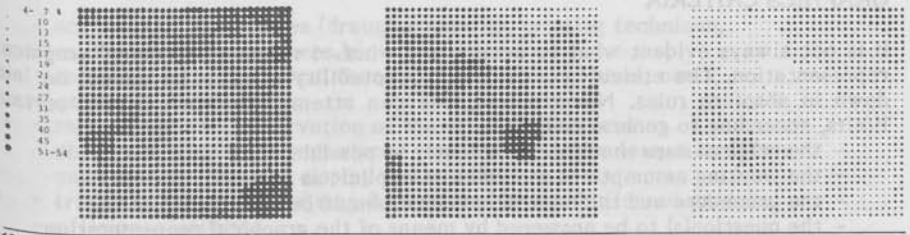


11.4 Indication of min. and max. values



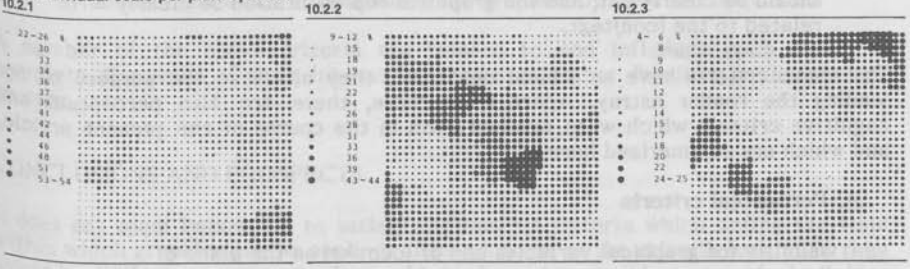
10.1 Common scale for all (three) plant species

10.1.1 10.1.2 10.1.3



10.2 Individual scale per plant species

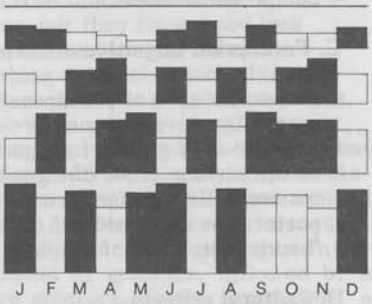
10.2.1 10.2.2 10.2.3



What is shown in Figure 10 for a map area can also be seen in dot matrices or multiple bar charts, as in Figure 11. The rows each represent a certain variable, the columns a unit of observation or registration. The chosen units of scale, per variable, are arbitrary. Figure 11.1 displays data which are not bound to a definite order, either horizontally or vertically, and which are rearranged. The reordering would be made more evident if quantities above or below a reference line were directly perceptible. In Figure 11.2, the shaded areas represent the above mean data per row.

Figure 11.3 is the result of two operations, induced by the wish to enhance marginality: first the zero lines have been replaced by minimum lines, then the maximum bar heights per row have been taken as a standard, with the other bars drawn in proportion. Figure 11.3 does not give an indication of absolute quantities: to find those, the reader would have to consult the corresponding datamatrix. Another solution is shown in Figure 11.4, where maximum and minimum quantities, per row, are indicated. The reader may then reconstruct the original data, but as the essential aim of the graphical representation is to show the nature of marginality per variable, the knowledge of absolute quantities is secondary. In Figure 12 we have a different case. First, the columns represent the months of the year, and as such do not have an arbitrary order. Secondly, the rows represent classes of only a single property (for example age classes). The sum of bar heights per column comes to 100% in each case. The black bars represent an enhancement of above mean values per row. Such graphical enhancements should, in any case, be accompanied by explanations, either in the legend or in the accompanying title and text.

12 Percentage matrix-chart with varying row height and enhancement of above-mean values



GRAPHICS CRITERIA

It is not always evident what to accept and what to reject in terms of graphical representation. The ethical or scientific acceptability of graphics cannot be laid down in absolute rules. Nevertheless, one can attempt to work within certain limits, according to general criteria:

- the original data should be as directly as possible;
- the working assumptions should be as explicit as possible;
- the procedure and the conventions used should be simply deduced;
- the question(s) to be answered by means of the graphical representation should be clearly put, and the graphical representation be clearly related to the (con)text.

The above criteria have an ethical character: they all serve the purpose of not leading the reader astray. Apart from these, there are also perceptual and cognitive criteria which were touched upon in the course of the present article, and which are summarized here:

A. **Perceptual criteria**

- visibility (of graphical variables and of location on the plane of display);
- immediacy of reception and response;
- clarity (avoiding perceptual confusion, e.g. optical illusions, flittering and other such effects);
- simplicity (no superfluous signs, decorations, colour for colour's sake, etc.);
- consistency and comparability (always use the same meaning for a sign in a given context, use visual properties consistently).

B. **Cognitive criteria**

- logical consistency (in order and in relations; no confusion of conceptual levels);
- accuracy and reliability of data;
- clarity (of classes, relations and properties);
- conciseness (no superfluous information);
- comprehensiveness.

C. **Perceptual-cognitive criteria**

- correspondence of perceptual and cognitive properties (in particular, levels of measurement and perceptual characteristics, expression of growth by logarithmic scales, constant proportionality of the surface areas of signs as an expression of quantities, etc.);
- memorability (design structure as an expression of knowledge structure, patterns as expression of quantities, etc.);
- 'heuristicity' (use of graphics as instruments of discovery).

D. **Cultural criteria**

- acceptability (if too new or too different, a graphic presentation runs a greater chance of being rejected);
- learnability;
- communicability (easy use of conventions);
- attractiveness.

E. **Technical/pragmatic criteria**

- accuracy of the Figures (draughtsmanship, printing technique, resolution of displays, etc.);
- technical and economical feasability (in terms of time, costs, expertise, hardware);
- reproducibility (preservation of the original graphical characteristics when reproduced - copies, enlargements, microfilms, and from one mode of display to the other);
- triability (possibility of step-by-step and bit-by-bit introduction);
- efficiency/utility;
- dynamics (possibility of reordering the graphics by hand or computer).

A number of the above criteria are related to and influence each other. A complete discussion of their interaction, however, is not feasible in the present context.

CONCLUSIONS AND PROSPECTS

It does not seem impossible to satisfy the various criteria which define the field within which graphics can be applied, although some of the criteria look as if they would be difficult to reconcile. It should, in short, be possible to produce graphics for decision support which are:

- concise without loss of relevant details;
- quickly comprehensible but not oversimplifying;
- geared to indentifying relationships and trends, but without loss of original data;
- rearrangeable at the services of the decision-maker, designer or analyst, to allow for new policy insights, but without too much effort, time, or intervention of specialist knowledge.

The problems lie rather on another level:

- graphicacy education: on the one side the familiarization with graphics principles, on the other side the production and marketing of learning material;
- software production for integrative and interactive microcomputer graphics use, superseding the spreadsheet generation, according to the criteria which have been listed here. It is true that a waterfall of software packages have reached the market, but they show a sad lack of graphicacy in their authors and buyers (and a love for vivid colours, three-dimensional impact and impressive flashing devices). Actractiveness seems to be the only prevailing criterion to have directed their development.

As Calvin Schmid underlines (14), it is only very recently that a climate has developed 'more conductive to the progressive development of graphic presentation in terms of higher standards, innovation, improved techniques, better trained specialist and wider usage and acceptability than it has been in several decades'. Graphics have known many ups and downs since the first medieval attempts. In our century, the twenties have seen a swift development of graphics, followed by a theoretical stagnation, as statistics became more abstract. At the same time, a small number of graphical stereotypes became more popular, in particular in the US. Triumphant American management has vulgarised them, and it is now the mummified knowledge of the twenties that is served to us with the help of advanced technology. But at the same time, the renewed interest for graphics creates the impulse Schmid is evoking. This could help to explain why the seminal

work by Jacques Bertin is now, at last, translated into English and Japanese (and a German translation in 1974) (2).

Bertin and his team have been developing his work further, not only by working out applications in various fields of science, but also by their computer graphics implementation (in particular J.D. Gronoff (15). At the technical University of Eindhoven, pilot studies are being done on a 'Graphical Pattern Processor' which uses interactive computer possibilities according to the principles and criteria which have been sketched here (16).

The potential of all this lies in realizing a symmetry between the large human talent for assimilating visual (and in particular graphic) information, and the skill for producing it quickly and efficiently. The computer could make us all fantastic graphical communicators, if the right means were offered. But then in the play of supply and demand, the demand side must make itself vocal. Instead of having to choose between programs developed on the basis of their selling appeal, decision-makers, designers, researchers, and other professional users should present the industry with their own brief. Such a brief cannot, as yet, be established on the basis of sound experimental research results within an integrating theory. There is still al lot of research to be done, but at least one could begin with respecting the principles of experienced and conscientious researchers.

NOTES

1) W.G.V. Balchin, A.M. Coleman: Graphicacy should be the Fourth Ace in the Pack, **The Cartographer**, 3 (1966) 23-28; W.G.V. Balchin, Graphicacy, **Geography**, 57 (1972), 185-195, as quoted in C.F. Schmid, **Statistical Graphics**, New York: John Wiley & Sons (1983) p.11.

2) J. Bertin,: **Sémiologie graphique: les diagrammes, les reseaux, les cartes**, Paris-La Haye: Mouton; Paris: Gauthier-Villars (1967 and 1973), german translation **Graphische Semiologie**, Berlin and New York: Walter de Gruyter, (1974), english translation **Semiology of Graphics**, Madison: University of Wisconsin Press (1984), translation in japanese (data unknown).
See also: J. Bertin, **La Graphique et le Traitement Graphique de l'Information**, Paris: Flammarion, (1977); english translation: **Graphics and graphic information processing**, Berlin and New York, Walter de Gruyter (1981).

3) Most of the examples are based upon illustrations of articles in the dutch journal for policy analysis "Beleidsanalyse", (see notes 5, 6, and 7).

4) For a review of current research, see H. Wainer, D. Thissen, Graphical Data Analysis, **Annual Review of Psychology**, 32, (1981), 191-241.

5) See H.L. Klaassen, J.W. Weehuizen, Een denkkader voor beleidsonderzoek, deel B, Evaluatiemethoden nader beschouwd, **Beleidsanalyse**, Vol. 10, nr. 3-4, (1981) p.13.

6) See F.J.P. Heuer, Over het uitvoeren van beleidsanalystische studies, **Beleidsanalyse** Vol. 9, nr. 1, (1980) p.11.

7) As illustrated in H.L. Klaassen, J.W. Weehuizen, Een denkkader voor beleidsonderzoek, deel B, Evaluatiemethoden nader beschouwd, p. 11.

8) See J.L.T. De Jong, Multi-Criteria-analyse: Een toepassing van de permutatiemethode van Jaquet-Lagrèze, **Beleidsanalyse**, Vol. 11, nr. 3, (1982), p.27-28.

9) See J. Bertin, **Graphics and graphic information processing**, Berlin and New York, Walter de Gruyter (1981), p. 34-36.

10) See J. Bertin, **Graphics and graphic information processing**, p. 5-6.

11) See books like Bertin (note 2) and C.F. Schmid (note 1) for the fundamentals.

12) See C.F. Schmid, **Statistical Graphics**, New York: John Wiley & Sons (1983) p. 65-68.

13) Stichting Bouwresearch, **Technologische en structurele ontwikkelingen in de Bouw**, Deventer/ Den Haag, Kluwer/Ten Hagen, (1980) p. 87.

14) See C.F. Schmid, op. cit. p. 5-6.

15) Author of programs, written and developed with Hewlett-Packard's 9845 T Minicomputer in HP Basic, Centre Pluridisciplinaire De La Vieille Clarité. 13002, Marseille, France.

16) Programs, written and developed by V. Tabery with Perkin Elmer 3220 Minicomputer, draughting for pilotstudies and this article by P.A.C. Rooijakkers and A.M.C. de Caluwe, Technical University of Eindhoven, P.O. Box 513, postvak 5, 5600 MB Eindhoven.

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

STRATEGIC CHOICE AND EVALUATION: SOME METHODOLOGICAL CONSIDERATIONS

P. Dello
Catholique University Leuven
Dept. of Social Geography

1. INTRODUCTION

One of the central themes of this workshop is the role of evaluation in strategic choice. In this paper we will approach this theme by focusing on the methods of evaluation used in strategic choice.

First, we will discuss the methods one finds in Hickling (1974), Hickling, Wilkin and Debreyne (1980) under the heading: 'Discrimination between Alternative Solutions'. We will highlight some characteristics of these methods, and demonstrate their consequences for the quality of the evaluation. The idea of field specific evaluation will be introduced as a means to activate the learning activity, which we consider to be one of the most important products of evaluation work.

Secondly, we will discuss an activity which, in strategic choice, is placed under the heading 'Structuring the Problem'. This is an exploration of the relations between options of different decision areas. We will argue that the option bars and uncertainty lines of the strategic choice toolbox are not sufficient for this work. We will demonstrate that this type of work has a strong evaluation dimension, and that a new method (the optionvector) can be introduced to make this characteristic more explicit. The principle of the method, and its advantages, will be stressed.

Finally, we will mention some elements which do not immediately relate to the formal evaluation activities of strategic choice; namely the decisive role played by scenarios and policies in introducing conditions for the formulation of the problem, when working with consistency matrixes and the DOT facility. Their importance for the evaluation work will be stressed.

Throughout the article, we will use some terms which are typical of strategic choice (decision area, option, option bar, solution).In order to keep the text readable for those who are not familiar with these terms, we will give a brief explanation of each, based on Hickling (1978 p. 470-474). A decision area can be described as any field of choice in which it is possible to formulate two or more mutually exclusive alternatives (:options). It is possible to identify incompatibilities between the options of different decision areas (formulating option bars). Although the option bar simplifies the relationship between options, it is a necessary device to manipulate the large number of possible combinations between options. Once the set of options in each decision area, and the incompatibilities between pairs of options, have been built up, it is a matter of straightforward logic to work out how many combinations of options are possible. A feasible solution is a combination of options on the basis of one option per decision area, taking into account all the incompatibilities expressed by the option bars.

In addition to a familiarity with these typical terms, the reader should also be

aware that we will implicitly work with the hierarchical structure strategic choice philosophy - concepts - methods as developed in Mastop, Faludi and Vermeulen (1981 p. 48-55). In this hierarchy, the concepts (uncertainty, cyclic and continuous process, choice situation, commitment package) function as a bridge between the philosophy and the methods. We will also refer to the concepts uncertainty, and cyclic and continuous process, both of which aim at activating learning processes.

2. DISCRIMINATION BETWEEN ALTERNATIVE SOLUTIONS

From now on, we will use the term instruction books to indicate three publications which cover the methods of strategic choice in a systematic way, namely Hickling (1974), Hickling, Hartman and Meester (1976) and Hickling, Wilkin and Debreyne (1980). The latter publications (1976 and 1980) are translations (respectively in Dutch and in French) of Hickling (1974).

The methods used in the three books are basically the same, although there is a period of six years between the publication of the first and the last one. In the intervening period, the main author (A. Hickling) has had several opportunities to work with the methods, and to test their reliability. As there are only minor differences between the three books, we can conclude that the methods concerned are seen as being useful and sufficiently operable within strategic choice.

In the instruction books, the methods are presented in a systematic way following the lines of the process of strategic choice. The process is built around four modes of working (shaping, generating, comparing, choosing), and the evaluation methods are, in the first place, grouped in the latter two modes. We will now focus on some of the characteristics of these methods in order to highlight some consequences which are, according to us, insufficiently stressed in the existing strategic choice literature.

The main activities which appear under the heading 'Discrimination between Alternative Solutions' in the instruction books can be summarized as follows.

1) The options are confronted with a set of cardinal (metric) criteria, and cardinal criterion scores are given (Table 1).

	Options		
	DA 1	DA 2	DA m
Criteria	$OP1.1 \ldots OP1.I_1$	$OP2.1 \ldots OP2.I_2$	$OPm.1 \ldots OPm.I_m$
1_c			
2_c			
.			
.			
J_c			

DA = Decision Area
OP = Option
$I_i \in \{2,3,\ldots\}$ with $i \in \{1,2,\ldots,m\}$

Table 1. **Scoring the options for metric criteria**

2) The cardinal criterion scores are transferred to the solutions which have been generated in a previous phase (Table 2).

Criteria	Solutions 1 2 ... n
1_c	
2_c	
.	
.	
.	
J_c	

Table 2. **The values of the solutions for the metric criteria**

3) The solutions are confronted with a set of ordinal criteria and rankings are given (Table 3).

Criteria	Solutions 1 2 ... n
1_o	
2_o	
.	
.	
.	
J'_o	

Table 3. **The values of the solutions for ordinal criteria**

4) All the criterion scores are transferred to rankings.

5) The dominancy principle is used for short-listing the list of solutions. by definition, a dominated solution is one which is inferior to another on the basis of at least one criterion, as well as not being superior on the basis of any other criterion.

2.1 The idea of field specific evaluation

Basically, the evaluation is not carried out at the level of the decision area, although this is perfectly possible. The formulation of criteria is not oriented towards the specific content of the decision areas, but occurs at a more general level. The quantitative criteria are scored on the level of the option, but for a great number of options, the quantitative criterion at hand is not relevant, and a zero will have to be used to show this. The qualitative criteria only come into play at the level of the solution. Their potential as an aid in gaining insight at the level of the decision area is completely ignored.

In the concept of strategic choice, the learning aspect is heavily stressed. By not evaluating at the level of the decision area, a whole range of learning chances are being lost. The formulation of field specific criteria forces the members of the working group to look at the decision object from different angles. Discussions about the potential impact of the different options of one decision area can be enhanced by working simultaneously with metric and with ordinal criteria at that level. As a result of this type of work, a feed back to the 'shaping' mode of the process is possible. Now options can be formulated, and others can be revised. Support for this activity is found in the concept of cyclic and continuous working.

So why are there no evaluation activities included at the level of the decision area?

One reason might be that in this way, a lot more information is made available, and that this surplus of information inhibits the progress of the working group. This argument holds in situations where information is manually processed. The microcomputer has now progressed to the point that powerful machines are available at reasonable prices. All the surplus information obtained at the level of the decision area can now be stored in a structured and systematic way. Since it is available at any time, in an operational form, it is no longer so overwhelming, and can be used without inhibiting the progress of the working group.

The next step, of course, is working with these field specific criteria. What can be done with them? Several possibilities exist. The degree to which the set of criteria is used depends upon the specific situation at hand. One way to proceed is to rank the options using the set of criteria, and to search for 'the good options' by simply working with the dominance principle. The other extreme is to score the cardinal criteria as well as possible on a cardinal scale, and to score the ordinal criteria on a 1 - 5, 1 - 7, 1 - 9 or 0 - 10 scale. Supplementary aids (methods) can be introduced to assist in this work (e.g. 0 - 1 pairwise comparison, Saaty pairwise comparison, etc.). A good survey of these aids (preference measurement methods) can be found in Voogd (1980, 1983).

We will not discuss this scoring and rating of criteria in detail, because if we did so we would have to deal with specific characteristics of a number of methods. This is beyond the scope of this article. It is clear, however, that an adaptive and flexible way of working is possible, thanks to the availability of these methods.

This discussion of the potential role of field specific criteria gives us the chance to move on to a second major consideration, introduced in the 'Discrimination between Alternatives'. The criteria (ordinal and metric) are not perceived to be of equal importance. A criterion expressing the housing needs in a neighbourhood can be of much more importance than one about the aesthetic nature of that neighbourhood. Everything depends on one's point of view. Different people (options) will have different ideas about the relative importance assigned to a set of criteria.

2.2 Some arguments against the idea of internal consensus in the working group

None of the methods used in the instruction books for strategic choice is designed to express different opinions within the working group. Once again, we believe that this situation is not realistic, and it contradicts the principles advocated via the strategic choice concepts. One obvious example of potential internal disagreement is the relative weight which can be attached to the criteria. In principle, this internal disagreement can occur in any strategic choice activity (formulation of option bars, scoring the criterion scores).

If one wants to activate a learning process within the group, these different opinions have to be made explicit. This explicit treatment of different opinions obviously creates a lot of surplus information. Thus, one really needs the necessary tools to store this information in an orderly way, and to work with it in an operational form. If these conditions are fulfilled, sensitivity analyses are very useful to activate a mutual learning process within the working group. We have found that the basic structure of the multicriteria evaluation methods (the evalua-

tion matrix and the priority matrix) can be useful in collecting information at the level of the decision area (tables 4 and 5).

Criteria	Options OP1.1...OP1.I_1	OP2.1...OP2.I_2	OPm.1...OPm.I_m
1_c			
2_c			
.			
.			
J_c			
1_o			
2_o			
.			
.			
J'_o			

Table 4. **The evaluation matrix**

Opinions	Criteria 1_c	2_c	...	J_c	1_o	2_o	...	J'_o
1								
2								
.								
.								
K								

Table 5. **The priority matrix**

In this structure, the internal disagreement can be expressed in the priority matrix. There, one can decide how important a particular criterion may (or should) be within that decision area. This activity can be seen as a first round. In a second round, a group discussion could be held in order to present the arguments underlying the scores of each member. Then, as in the Delphi process, the opportunity could be given to revise the personal scores (third round). In a fourth round, a multicriteria method could be used to find out how the options get ranked on the basis of the different opinions (the preference scores table 6).

Opinions	Options OP1.1...OP1.I_1	OP2.1...OP2.I_2	OPm.1...OPm.I_m
1			
2			
.			
.			
K			

Table 6. **The preference (appraisal) matrix**

These results can be the basis of a new group discussion, taking into account that the preferences are not only influenced by the priorities, inconceivable. In the literature of multicriteria analysis, the criterion scores are supposed to be deter-

mined by experts, and are thus beyond discussion (Voogd 1980a, 1980b, 1981a, 1981b, 1981c, 1982, 1983). In a learning situation, however, there is no reason why the criterion scores should not be the subject of internal disagreement. Several runs could be undertaken to see when (and under which circumstances) a shift in option preference occurs as a consequence of a shift in the criterion scores or in the priorities. In this way, one can explore the limits and the consequences of the different ideas within the working group.

Such learning processes can be organised in several ways, and the depth of this analysis can be adapted to the importance of the decision area. It is, of course, not necessary nor realistic to explore all the decision areas by means of field specific criteria. We do believe, however, that for crucial decision areas this work can be very rewarding.

We have undertaken a review of the strategic choice literature in order to find out whether this idea of field specific evaluation has been used before in the practice of strategic choice. We have analysed the studies written in capital letters in table 7 (based on Mastop and Van Rosmalen 1981 p. 11). Table 7 groups the strategic choice studies with an emphasis on practical work (applications). We have discovered that the formal 'Discriminition between Alternative Solutions', as described in the instruction books, is only applied in a few cases (Mastop and Dekker 1979; Mastop, Faludi and Dekker 1979; Dekker and Mastop 1979). Another example, not mentioned in table 7, but strictly developed as described in the instruction books, is that of Arnhem (1978). In none of these studies, nor in the other ones where the evaluation work is carried out in other ways, does field specific evaluation occur.

3. RELATIONS BETWEEN OPTIONS OF DIFFERENT DECISION AREAS

One of the concepts of strategic choice, 'Choice Situation', is built around the idea that choices about different matters should be linked with each other in order to see how they influence each other. This idea has been brought into practice by means of AIDA (Analysis of Interconnected Decision Areas), where special attention is given to decision areas which are interconnected. Interconnected means, in the first place, that one can not make a decision in one area without refering to the decision to be made in another area.

Once the interconnectedness is known at the level of the decision area in AIDA, one proceeds to the analysis of interconnectedness at the level of the options. In AIDA terms, this is called the formulation of option bars and of uncertainty lines. An option bar indicates that a particular combination of options is impossible. The rationale behind the bar is mainly of a technical nature. An uncertainty line indicates that a certain combination of options is in some way conditional. The conditional nature may be caused by several factors. These factors can be summarised and mapped on the well known radial chart of the strategic choice methodology (Figure 1).

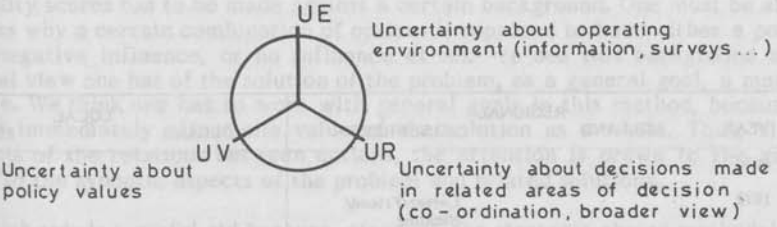


Figure 1. **The radial chart**

We will now focus on the UV uncertainty (uncertainty about policy values), and relate it to the uncertainty line. It often depends on personal values whether or not a specific combination of options is felt to be impossible. Moreover, it is often very unrealistic to force a 0 - 1 statement about a combination of options. There exists a whole range of desirability between the extremes of accepting the combination or rejecting it. In the strategic choice literature, this fact has been given little attention. In only three of the studies we analysed (see Table 7), has the principle of a detailed analysis of the relations between options been raised (Bather, Williams and Sutton, 1976; Sutton, Hickling and Friend, 1977 and Amos c.s., 1977). We know of one other study where that topic is introduced via the question of additivity. "It is assumed that if policy A has an effect of 9 on problem X upon which policy D has an effect of 13, then if A and D are in a strategy they have a combined effect of 22 upon X. This is not necessarily the case as policies when combined could have a detrimental effect upon problems. However, it is an assumption which has to be made in the absence of a detailed policy compatibility analysis".

This idea of policy compatibility analysis has been made concrete to a certain extent in Sutton, Hickling and Friend (1977) and in Amos (1977). In the former study a -3 to +3 scale is proposed, with -3 expressing a very negative interference (almost equal to an option bar), 0 the neutral situation, i.e. no mutual interference, and +3 a positive mutual reinforcement. A similar proposition has been made in Amos (1977) but this time with a 0 to 4 scale. Both studies have the advantage of making the nature (positive or negative) of the mutual influence between options explicit by means of scores, but neither of them solves the problem of additivity. If, for instance, in a particular situation there are five positive combinations of options (+3) and five negative combinations (-3), the overall score would be zero. If, in another situation, there were nine combinations with score zero and one with score -1, the overall score would be -1, and one might draw the false conclusion that, in the second case, the internal coherence is worse than in the first case. In other words, in Sutton, Friend and Hickling (1977) and in Amos (1977), a lot of interesting information is being collected, but no device is given to manipulate it in operational terms.

We have developed a method (the **option vector**; Dello and Gheldof, 1983) to make this explicit information about the relation between options operational. The basic idea is that the new information is linked with the preference scores of the options involved. In this way a quadratic function has been developed (thus no linear summing). With the linkage, we propose that the quality of a solution is no longer only determined by means of the preference scores of the options, but also by a factor which we might call the internal coherence. This exploration of the quality of the relations between options (determining the compatibility scores), and of making them operational, again offers chances to activate the learning

YEAR	REGIONAL		LOCAL	
	HOLLAND	ELSEWHERE	HOLLAND	ELSEWHERE
1970				Logimp
1972		Carter/Friend/ Hickling		
1973				Bunker/Frie
1974				Bunker
1975		BATHER/WILLIAMS/ SUTTON Ferguson		Cardinal/ Fellows/Dav
1976	BARDIE/ DEKKER/ GORTER (A&B)	Environmental Friend/Sutton	MEESTER/ HARTMAN	
1977	Bartelds/ Cats/ Rubingk BARDIE e.a. VINK	Alternative SUTTON/HICKLING/ FRIEND		
1978	HORST/ JANSEN/KLOK VAN HUUT DEKKER/MASTOP/ VERDUIN	Hickling/ Luckman/	DEKKER/FALUDI/ MASTOP DEKKER/MASTOP	
1979	Faludi/ Dekker/ Mastop MASTOP/DEKKER MASTOP/FALUDI/ DEKKER VOOGD	HICKLING/FRIEND/ LUCKMAN	DEKKER/ MASTOP De Jong Terv.	Bundgard Jorgensen
1980	Vos		Bestemmings- plan Handleiding	
1981	MASTOP/FALUDI VERMEULEN VERMEULEN/DEKKER/ MASTOP		Structuurplan	

Table 7. **An overview of the strategic choice literature with an emphasis on practical work**

process within the working group. It is clear that the determination of the compatibility scores has to be made against a certain background. One must be able to express why a certain combination of options is supposed to have either a positive or a negative influence, or no influence at all. We see this background as the general view one has of the solution of the problem, as a general goal, a main objective. We think one has to work with general goals in this method, because the scores immediately affect the value of the solution as a whole. Thus, via the analysis of the relations between options, the attention is drawn to the general view, to the synoptic aspects of the problem and related solutions.

We think this is a useful aid because, after all, the strategic choice methodology is dominated by an analytic way of working (breaking down the problem into manageable pieces ...). The option vector offers the chance to start a discussion about general goals, and this is, according to us, a phenomenon worth testing in the practice of strategic choice.

We see no practical restrictions on its use. The principle of the method can be explained in a simple (even graphic) way, and can be separated completely from its more complicated, technical elaboration when one proceeds to a n-dimensional situation. The input one needs consists of explicit statements, comparable with the ones needed, for instance, for scoring the options. Last but not least, the option vector fills the gap in the strategic choice methods mentioned by Bather c.s. (1976), Sutton c.s. (1977), Amos c.s. (1977) and Hayton (1977).

The principle of the option vector can easily be demonstrated in the case of two options (O_1 and O_2 in Figure 2).

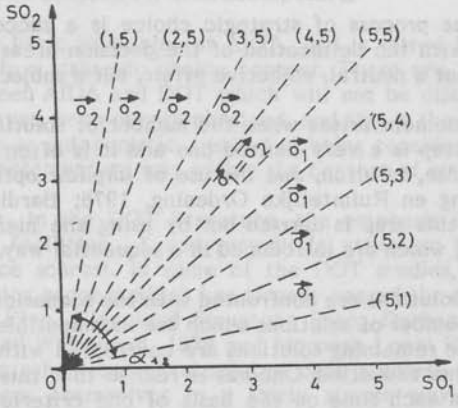


Figure 2. **The principle of the option vector in the case of two options**

Two perpendicular axes (so_1 and so_2) are drawn, and the scale which is used for the compatibility scores is indicated on both axes. To simplify the calculations, we have used a 1-5 scale (1 = very negative, 2, 3 = neutral, 4, 5 = very positive). On so_1, the relation of O_1 with itself (obviously +5), and with all the other options, in this case only option O_2 (with a score of e.g. 3), is drawn. This results in a point with coordinates (5,3). The connection of (0,0) with (5,3) gives the orientation of the option vector O_1. The length of the option vector will be given by the preference score of option O_1. Figure 2 shows the different orientations of the option vectors which are possible with a 1-5 scale. The maximum angle between

two option vectors (12 in Figure 2) corresponds with the situation (1,5 - 5,1), which means that the relationship becomes more positive (2,5 ; 5,2) - (3,5 ; 5,3) - (4,5 ; 5,4), the angle between the option vectors diminishes. In the case of a positive mutual influence (5,5 ; 5,5), the angle does not exist.

This angle is the new element introduced in this method. Linear summing of the preference scores of the options will, from now on, occur in only one situation (maximum mutual reinforcement). In all the other cases, the width of the angle between the option vectors will determine the factor with which the sum is reduced, because the sum is given by the principle of the parallelogram of forces.

We believe that the principle of the option vector is fairly simple, and this is an important argument for adding the option vector to the toolbox of strategic choice. The other argument for doing so is the learning effect it can induce. First of all, it offers the opportunity to organise a discussion at the general level of the problem, and secondly, it helps to indicate solutions which are internally consistent. This new information can become very valuable in the choosing phase of strategic choice.

Instead of working with one opinion, one could also apply the compatibility scores of one person to a set of solutions preferred by another person. If the value of these solutions diminishes sharply, this would mean that both persons disagree. This fact could be the starting point for a discussion about this disagreement.

4. SOME REMARKS ON LESS EXPLICIT EVALUATION MOMENTS IN STRATEGIC CHOICE

One can say that the process of strategic choice is a succession of evaluation moments. It starts with the delineation of the decision areas and of the options, something which is not a neutral, objective affair, but a subjective decision.

Another evaluation moment arises when the number of solutions is diminished. In several studies this step is a very limited one and it is often based on such principles as common sense, intuiton, and the use of implicit option bars (e.g. Ministerie Volkshuisvesting en Ruimtelijke Ordening, 1976; Bardie c.s., 1977a,b). In some other studies, this step is carried out by using the hierarchy of scenarios, policies, and actions, which are introduced in a sequential way.

In a first round, the solutions are confronted with the scenarios, and this results in an elimination of a number of solutions which are incompatible with the scenarios. In a second round, the remaining solutions are confronted with the policies, which in turn causes a further reduction. One has to realise that this procedure causes a sequential evaluation each time on the basis of one criterion, and that the sequence is decisive for the weight attached to the scenarios and the policies. Yet this dominant role of the scenarios is not explicitly underlined in the studies mentioned above.

A last item which we want to mention is the Decision Optimising Technique (DOT) and its relation to evaluation. There are several articles on DOT, and these reflect the evolution of the method between 1975 and 1982 (Table 8).

Classification by content	
1. Strictly technical	1. O.75
2. Emphasis on theoretical aspects (without example)	2. O.76 - W.79 - O,&W.80a - O.&W.80b
3. Emphasis on practical aspects (usefulness and applicability with example)	3. O.&W.77 - O.&W.78 - O.&W.79 - O.&W.82
Classification by type	4. O.76 - O.&W.77 - O.&W.78
4. DOT 1	5. O.&W.79 - W.79 - O.&W.80a - O.&W.80b
5. DOT 2	6. O.&W.82
6. MicroDOT	

O.76= Openshaw (1976); W.= Whitehead; O.&W.= Openshaw and Whitehead.

Table 8. **A classification of the DOT literature**

The User's Guide of DOT 1 is strictly technical. DOT 1 is a method to solve and optimize the combinatorial problem simultaneously. DOT 2 is an extention of DOT 1. In addition to the DOT 1 activities, a search process can be organized to reach a compromise solution between several opinions. MicroDOT is basically a DOT 2 version which has been adapted for microcomputers.

DOT is built on the basic principles of AIDA, and is suited for the type of work that has to be done in a strategic choice context. There are, of course, a number of differences between AIDA and DOT which will not be discussed here. The differences have, however, been overemphasized, and this is the reason why DOT and AIDA are often seen as antagonistic, instead of being considered as methods which can be used in a complementary way within the context of strategic choice.

We have found that, in the DOT literature, the emphasis is on methodological aspects (capacities, facilities ...), and very little attention is given to the DOT input (the preference scores). In some of the DOT studies, the input is copied blindly from a planning process which has already passed the evaluation phase (e.g. Cumbria Structure Plan, Cleveland Structure Plan, Durham Country Structure Plan in Openshaw and Whitehead, 1979 and Morpeth Local Plan in Openshaw and Whitehead, 1977). In other studies (simulated case studies), the authors produce the preference scores themselves in a fairly simple way, without using any specific preference measurement method to give some depth to the input (e.g. Blankshire Structure Plan in Openshaw and Whitehead, 1978 and Red Lane Action Area Plan in Openshaw and Whitehead, 1982). This fact is relatively surprising, because DOT offers some facilities which prove the author's interest in the evaluation dimension of the planning work.

We would like to mention two facilities which we have found to be useful, and which are not available in the strategic choice toolbox. The first one is the so called random facility, which distributes the preference scores in a random way within limits imposed by the user. These limits (e.g. 10%, 20%, 30%) can stand for a degree of uncertainty which a person can have about his preference scores. A number of DOT runs can be undertaken for each degree of uncertainty, to test the

extent to which the 'optimal solutions' change as a function of the uncertainty introduced. This facility is useful for getting an idea about the robustness of a solution in a short period of time.

The second facility we want to highlight is the possibility of introducing all sorts of conditions in the DOT problem formulation. This is done by means of equations (smaller than, bigger than, equal to). In this way, a whole range of hypothetical situations can be tested, and this is, of course, very helpful in the final phase of choosing a solution.

5. CONCLUSION

In this article, we have explored some difficulties which arise when working with the classic tools of the strategic choice toolbox.

We have shown that the structure of the multicriteria evaluation methods can function as a basis for collecting explicit information which can be made available at the level of the decision area. This proposition can only be made operational if one accepts that evaluation work, at the level of the decision area, is useful, and if one is willing to abandon the idea of implicit consensus within the planning (working) group.

A second major conclusion is that 0 - 1 statements concerning the relations between options of different decision areas are insufficient. A more subtle exploration of these relations can be very useful. This is an opportunity to work with the problem and related solutions in a synthesizing manner. This is something new in an approach which is dominated by analytical methods.

A third conclusion is that almost every activity in the strategic choice process has an evaluative dimension. We found that in practical working conditions, one is not always aware of this situation. This can lead to implicit evaluations.

Our final conclusion is that in DOT studies, no specific methods (preference measurement methods) are used to produce the basic input (the preference scores). Thus, the ideas of field specific evaluation, and of detailed policy compatibility analysis, also apply to DOT.

6. REFERENCES

Amos S., Alternative Strategies for Merseyside: A First Year Project Report, Department of Civic Design, Liverpool University (1977) (unpublished)

Arnhem, Structuurplan Arnhem, Keuzenota en Bijlagen, Municipality of Arnhem, The Netherlands (1978)

Bardie R., Dekker F., Gorter L., De strategische keuzebenadering: een toepassing, Oostrand Onderzoek, Rapport 5a, University of Utrecht (1976)

Bardie R., Dekker F., Gorter L., De strategische keuzebenadering: een toepassing, Oostrand onderzoek, Rapport 5b, University of Utrecht (1976)

Bather N.J., Williams C.M. and Sutton A., Strategic Choice in Practice: The West Berkshire Structure Plan Experience, Geographical Papers, No. 50, University of Reading (1976)

Dekker F. and Mastop H., Strategische keuze; een algemene planningsmethode ?, **Intermediair**, nr. 34 (1978) 47-59

Dekker F., Faludi A. and Mastop H., De strategische keuzeplanning en streekplanning. Handleiding voor de gesimuleerde case-study, Planologisch en Demografisch Instituut, Universiteit van Amsterdam (1978)

Dekker F., Mastop H., Verduin G., A Multi-Level Application of Strategic Choice at the Sub-Regional Level, **Town Planning Review**, Vol. 49, No. 2 (1978) 149-162

Dekker F. and Mastop H. (1979), Strategic Choice: an Application in Dutch Planning Practice, **Planning Outlook**, Vol. 22, No. 3 (1979) 87-96

Dello P. and Gheldof A., Het gebruik van multicriteria evaluatiemethoden in het aardrijkskunde-onderwijs, **De Aardrijkskunde**, (1983) 297-313

Hayton K., The use of a Decision Optimising Technique for Structure Plan Generation in Tyne and Wear, **Planning Outlook**, Vol. 22, No. 3 (1979) 97-101

Hickling A., **Managing Decisions.** The Strategic Choice Approach, Mantec, Rugby (1974)

Hickling A., Hartman R. and Meester J., **Werken met strategische keuze**, Samson, Alphen aan den Rijn (1976)

Hickling A., Aida and the Levels of Choice in Structure Plans, **Town Planning Review**, Vol. 49, No. 4 (1978) 459-475

Hickling A., Friend J. and Luckman J., The Development System and Investment Programmes, Final Report to DOE, COOR/19R, Coventry (1979)

Hickling A., Wilkin L. and Debreyne F., Technologie de la decision complexe. Des aides pour l'elaboration des choix strategiques, Universite Libre de Bruxelles (1980)

Mastop J.M. and Dekker F., Het streekplan Utrecht-Oost: een gesimuleerde case-study, Planologisch en Demografisch Instituur, Universiteit van Amsterdam (1979)

Mastop J.M., Faludi A. and Dekker F., De strategische keuzebenadering en streekplanning, Planologisch en Demografisch Instituut, Universiteit van Amsterdam (1979)

Mastop J.M., Faludi A. and Vermeulen A.H., Strategische keuze en streekplanning. E: Eindverslag, Verkenningen in Planologie en Demografie 22E, Universiteit van Amsterdam (1981)

Mastop J.M. and Van Rosmalen E., Strategische keuze en streekplanning. F: Verkenningen in Planologie en Demografie 22F, Universiteit van Amsterdam (1981)

Meester J. and Hartman R., Strategische keuze: een alternatieve benadering van planning, **Planning, methodiek en toepassing**, Nr. 1 (1976)

Ministerie van Volkshuisvesting en Ruimtelijke Ordening, Studierapporten Rijksplanologische Dienst Nr. 7, Beleidsalternatieven ten behoeve van de Verstedelijkingsnota, Den Haag (1976)

Openshaw S., FORTRAN IV Program and User Manual for the Decision Optimising Technique for Planners (DOT 1), Newcastle University, Department of Town and Country Planning (1975)

Openshaw S., An Alternative Approach to Structure Planning: The Structure Plan Decision Making Model (SPDM), **Planning Outlook**, Vol. 17 (1976) 10-26

Openshaw S. and Whitehead P., Decision Making in Local Plans, the DOT Methodology and a Case Study, **Town Planning Review**, Vol. 48, No. 1 (1977) 29-41

Openshaw S. and Whitehead P., Structure Planning using a Decision Optimising Technique, **Town Planning Review**, Vol. 48, No. 4 (1978) 486-501

Openshaw S. and Whithead P., The Generation and Evaluation of Structure Plans using a Decision Optimising Technique (DOT), in: Gullen I.G. (Ed.), **Analysis and Decision in Regional Planning**, Pion, London (1979) 178-202

Openshaw S. and Whitehead P., A Conceptual Framework and Computer based Methodology for Continuous Adaptive Planning, PTRC Summer Annual Meeting, 7-10 July 1980, University of Warwick, Planning and Transport Research and Computation Company Ltd., London (1980a)

Openshaw S. and Whitehead P., A Rejoinder to Willis and Thomson, **Town Planning Review**, Vol. 51, No. 1 (1980b) 71-75

Openshaw S. and Whitehead P., A Monte Carlo Simulation Approach to Solving Optimisation Problems Related to Plan Generation and Evaluation (microDOT), Paper presented at 22nd European Congress of the Regional Science Association, University of Groningen (1982)

Van der Horst J., Jansen C. and Klok I., Samenhangend ruimtelijk beleid? Een onderzoek naar de toepasbaarheid van de strategische keuzebenadering op provinciaal niveau, Oostrand Onderzoek, rapport 9, University of Utrecht (1978)

Van Huut H., Flexibiliteit: afschuiven van onzekerheden ?, **Stedebouw en Volkshuisvesting** 9 (1979) 390-399

Vermeulen A.H. Dekker F. and Mastop J.M., Strategische keuze en streekplanning D: Dommeldal; methodiek en planvorming, Verkenningen in Planologie en Demografie, 22D, Universiteit van Amsterdam (1981)

Vink H., Terugblik op een regionaal planningsproces: mogelijkheden en problemen bij de toepassing van doelstellingen en de strategische keuzebenadering in de prakijk, **Planning; methodiek en toepassing**, Nr. 3 (1977) 2-14

Voogd H., Het gebruik van multicriteria-evaluatie in de ruimtelijke planning, Planologisch Studiecentrum TNO, Delft (1980a)

Voogd H., Uncertainty and Multicriteria Evaluation. Planologisch Memorandum 80-8, Delft University of Technology (1980b)

Voogd H., Multicriteria Evaluation with Mixed Qualitative - Quantitative Data, Planologisch Memorandum, 81-7, Delft University of Technology (1981a)

Voogd H., Monitoring of Environmental Plans: an Exploration, In: Voogd H. (ed.), **Strategic Planning in a Dynamic Society,** Delftsche Uitgevers Maatschappij B.V., Delft, (1981b) 233-242

Voogd H., Qualitative Multicriteria Evaluation Methods for Development Planning, **Canadian Journal of Regional Science,** IV (1981c) 73-87

Voogd H., Multicriteria Evaluation with Mixed Qualitative and Quantitative Data, **Environment and Planning B,** Vol. 9 (1982) 221-236

Voogd H.,**Multicriteria Evaluation for Urban and Regional Planning,** Pion, London (1983)

Whitehead P., The Development of Decisionmaking Techniques for Planning, **Planning Outlook,** Vol. 22 (1979) 33-38

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EVALUATION IS A FIVE-FINGER EXERCISE

A. Hickling
Allen Hickling and Associates
Long Itchington, Rugby

INTRODUCTION

The conventional view of the rational process of decision-making and planning is one of a simple linear sequence of activities, one of which - somewhere towards the end - is evaluation. Unfortunately, attempts at following it consistently fail to confine the evaluation activity to only the one phase. Decision-makers and planners in practice make judgements which are often subconscious and intuitive, all the time, throughout the process (Lichfield **et al.**, 1975). That these judgements are evaluative in nature is recognised by frequent use of the phrase "value judgement" in their description.

The implication is seen to be that the practitioners, unable to follow the rational process, are acting irrationally. However, as is usually the case when something appears irrational, such a view indicates only that the underlying rationale is not understood.

A more helpful conclusion is that evaluation is essential throughout the process, and that it is impossible to make progress without making judgements all the time. Such evaluation may be as simple as an intuitive assessment that something is nice or nasty, or good or bad (Bentham, 1967). On the other hand, it may be a carefully considered opinion based on a much more sophisticated and elaborate analysis. In any case, there is much to be gained from a better understanding of the different evaluation styles, and the moments at which they occur.

This paper is written to this end.

THE DECISION-MAKING/PLANNING PROCESS

This view of evaluation as a continuous thread woven into the process at all stages implies that there must be many evaluative moments. So many, that it may well be impossible to predict when and in which order, if any, they are going to occur. In fact, it is most likely that the order will vary from project to project, and from time to time.

Building on earlier thoughts on the subject*, they are likely to vary in character according to the stage in the process at which they occur. And it is this which is the clue as to where to begin sorting them out. It should be possible to identify the different types, or styles of evaluation, related to the different activities in a decision-making or planning process.

In order to do so, it is necessary to adopt some form of model of the decision-making and planning process. And, as has already been pointed out, these are usually linear in form. But this new view of evaluation implies frequent skipping forward (or back) in order to accommodate the evaluation moments as they occur. And then some form of recycling is necessary in order to feed back the results of that evaluation.

In this light, the most helpful model currently in use in the Strategic Choice

process (Van de Graaf, 1985), which is cyclic rather than linear (Hickling, 1982). In the particular form used here, six activities (or modes of work) are identified:

scanning: which is no more that purposively observing the environment (in its broadest sense) in order to pick up information about the decisions to be made;

shaping: which involves the identification of those matters which are important, and those which are not, in establishing an orientation and focus for the work;

designing: which is the creation and development of alternative ways of dealing with the issues which are the current cause for concern;

comparing: which is the type of work most often associated with evaluation, and in which the differences between alternatives are explored;

choosing: which is generally thought to be the activity around which all the other decision-making activities are organised - usually located at the end of linear models of the process;

doing: which, sometimes coming under the general rubric of 'implementation', is often considered to be separate from the process, but which is an essential linking element in a cyclic process.

The activities could have been presented in any order because no sequence or priority is prescribed. However, in order to aid the transition from a linear view of the process, a sequence similar to that normally used has been adopted.

In fact, this does have another advantage in that it is possible on the same diagram (see Figure 1) to identify the 'locus' of each activity. For example, the scanning and doing modes of work must obviously take place in the environment or "Field of Operations". On the other hand, the designing and comparing modes are more 'back-room' activities which can be seen to take place in what might be called the "Technical Domain". Between the two areas thus identified lies a middle ground in which the relatively value-laden activities of shaping and choosing take place - sometimes called the "Political Arena".

THE "FIVE-FINGER MODEL" OF EVALUATION

In such a cyclic process, each mode of work or activity can be seen to have a product which makes a contribution to the product of the whole process - the decision which is reached or the plan which is made. In this case, it seems reasonable to assume that such a contribution is not in some way pre-ordained - there are alternatives to be considered.

Therefore there is an implication that choices have to be made about what these products should be - that, in fact, the work in each mode is a microcosm of the overall choice process. Thus, as the model just described is a model of that choice process, it must apply also to each activity.

For example, the product of the shaping mode is an orientation, or focus, for the work. Obviously there is not only one possible orientation. Alternatives will have been defined (designing in shaping), before a choice can be made between them

(choosing in shaping). Hopefully, the choice of an appropriate orientation will only be made in the light of an examination of the differences between the alternatives (comparing in shaping). And so on.

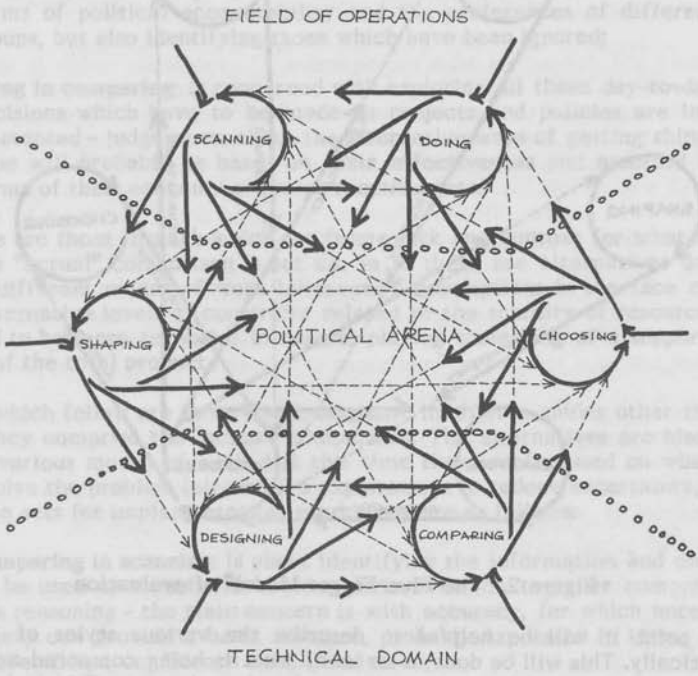


Figure 1. **The Cyclic Decision Process**

In practice, the effect is one of 'looping out' from the primary mode of work of the moment into others and back again - something which often happens quite quickly, and in many cases sub-consciously. And graphically, as there are six activities in this model of the process, the loops can be seen to form five fingers spreading out from each mode of work - one to each of the others - which is why it has been named "The Five-Finger Model".

All six activities are viewed as being pursued concurrently, although for any one person at any one time some are likely to appear to be of more significance than others. The part of the process being 'spotlighted' in this way will change over the passage of time, and according to who is viewing it (i.e. who is managing the 'spotlight'). As this paper is about evaluation, it is obvious that the 'spotlight' must be centred on the comparing mode of work. Therefore, from here on, this paper will be about only that activity and its five 'fingers', which are brought into play at different moments of the process - moments now identified in terms of which activity is in the 'spotlight' at the time.

THE DIFFERENT STYLES OF EVALUATION

The five "fingers" radiating from the comparing mode can be seen as different styles of evaluation as can those from all the other modes directed towards the comparing mode (see Figure 2). Thus it would appear that there are then different styles, none of which are normally differentiated in practice all of them referred to under the general rubric of evaluation.

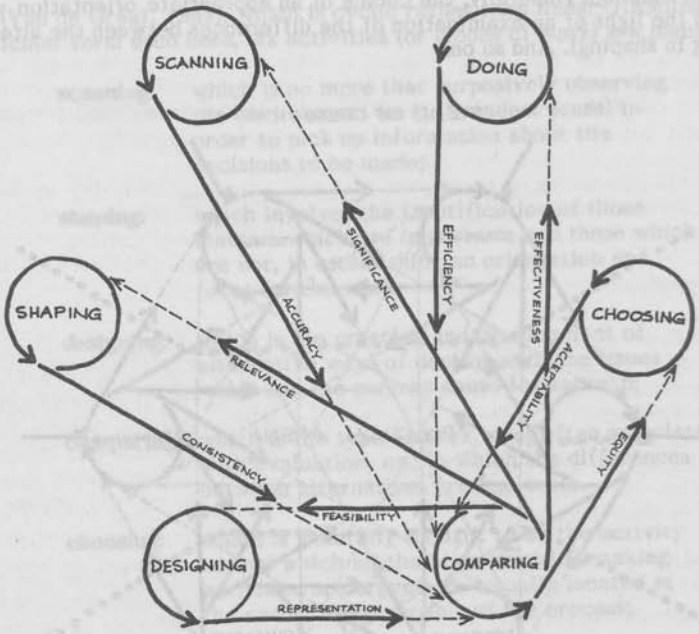


Figure 2. **"The Five-Finger Model" of evaluation**

At this point it will be helpful to describe the various styles of evaluation systematically. This will be done in terms of what is being compared and the sort of criteria which would be used in that style of evaluation. This is intended to capture the essence of each style. Further, they are grouped into two sets of five - those originating in the comparing mode, and those originating in the other modes - each set representing a somewhat different aspect of evaluation.

Starting with those originating in the comparing mode, each "finger" is now described in turn:

scanning in comparing: is about looking at alternative sets of explorations to reduce uncertainty (surveys, research, participation, consultation, etc.) aimed at helping future decision-making - judgements about which type of exploration, and which components of the environment to explore, are likely to be based on their **significance** and the degree to which current decisions are sensitive to them;

shaping in comparing: is the task of evaluating alternative foci and directions for the work, usually expressed in terms of clusters of choices and potential orientations - they are assessed according to their **relevance** at the moment by identifying which decisions are urgent, which will have a significant impact, and which are most closely inter-related;

designing in comparing: is concerned with sorting out alternative sets of possible solutions which cover the range of realistic ways forward - the most important consideration is that they are **feasible**, usually in relation to scarce resources, but also in terms of their compatibility with established policy;

choosing in comparing: is about discriminating between alternative policies and proposals designed to provide the basis for progress in the future - the selection of which to adopt will tend to be based on considerations of **equity** (see Miller, 1985), probably assessed in terms of political accountability and the preferences of different groups, but also identifying those which have been ignored;

doing in comparing: is concerned with exploring all those day-to-day decisions which have to be made as projects and policies are implemented - judgements about the alternative ways of getting things done will probably be based on their **effectiveness** and assessed in terms of their economy and time requirements.

These five are those through which the framework and context for what might be called the "actual" comparison is set up. In it there are alternatives associated with the different modes of work (alternative assumptions in the face of uncertainty, alternative levels of constraint related to the scarcity of resources, etc.). They tend to be more technical in nature, playing something of a supporting role in terms of the total product.

The five which follow are those which originate in all five modes other than comparing. They comprise the 'actual' comparison. The alternatives are also derived from the various modes of work, but this time they are focussed on what can be done to solve the problem (alternative explorations to reduce uncertainty, alternative action sets for implementation, etc.). They are as follows:

comparing in scanning: is about identifying the information and data to be used as a basis for the comparisons which together comprise the reasoning - the main concern is with **accuracy**, for which uncertainty can provide a basic measure, probably expressed in terms of probabilities, confidence limits, and so on;

comparing in shaping: is about defining the lines along which the comparisons are to be made - they should be **consistent** with the working focus, and likely to be defined in terms of which criteria to use, and which objectives have to be considered;

comparing in designing: is concerned with identifying a range of alternative solutions as an essential part of the framework within which the comparisons have to be made - the prime consideration is how well the alternatives **represent** the richness of choice available, without there being so many that informative comparison is impossible;

comparing in choosing: is about narrowing down the range of choice by identifying those which are less **acceptable** in a broadly political sense - probably expressed in terms of how effects should be traded-off one against another; or as limits below which achievement should not fall;

comparing in doing: is about limiting the range within which comparisons have to be made, by identifying and applying constraints imposed by the need for **efficiency** in the use of scarce resources - probably expressed in terms of budgets, time-scales, manpower, equipment and materials.

But all this refers to evaluation associated with decisions about the substantive problems under consideration. There is a whole range of other decisions to be made - choices about how that substantive decision-making process should be managed - so-called methodological choices.

METHODOLOGICAL CHOICES IN EVALUATION

In this paper the concern is with management of evaluation which is centred on the comparing activity. Choices have to be made continually about which concepts, frameworks and techniques should be used, when - in which sequence - in what combinations. Choices have to be made about how long they should be pursued and how intensively - and in what depth, or breadth. Choices have to be made about who should be doing the work - and how they should interact, with whom - and when - or not. Choices have to be made about the form in which the evaluation findings should be communicated. And so on.

In terms of "The Five-Finger Model" there are choices about how the fingers should be brought into play, which ones, when, and in which combinations and sequences. The different styles of evaluation are used according to the circumstances just as a pianist employs his or her fingers in playing the piano.

Any keyboard musician knows that 'five-finger' exercises are an essential basis to a consistently good performance. As the name implies, they involve the development of skills in the co-ordination of all five fingers in order to provide pleasing sounds. (Note that in piano-playing the concept of 'fingering' includes the thumb - but 'thumbing' is probably something else again, closely associated with noses and/or pages.)

The analogy is particularly apt when the pianist is improvising an exercise without the aid of a score. At any specific moment he or she chooses which finger or combination of fingers - or, indeed, which sequence of fingers or combination of fingers - should be applied to which note or notes. All five fingers will be used at different times and at different places on the key-board.

The application of the fingers to the keys may vary in strength according to the effect desired, and each configuration (confingeration?) will vary according to that which precedes it, and that which is intended next. There are usually alternative fingerings which could be used in order to achieve the same result. There is also, looking at the slightly longer term, the choice of where on the keyboard to play - and looking more widely - the choice of which instrument on which to play (piano, organ, clavichord, harpsichord, synthesiser, etc).

source: Radio Times (1985)

The point is that, as circumstances vary, the keyboard artist uses his or her resources selectively according to his or her capabilities and the effect desired. So it is with evaluation. The five styles of evaluation have to be used selectively - different stages of analysis of different types of problems involving different sets of people producing different evaluative moments in any decision-making or planning process.

And, where there are choices, the Strategic Choice model of the process again

applies. Thus the same modes of work come into operation:

scanning: for the future needs of and for evaluation, as well as moments of evaluation as they occur;

shaping: the focus for evaluation, and any significant limitations or requirements to be taken into account;

designing: the alternative ways in which the evaluation could be carried out, including who might be involved, and when, as well as the techniques which should be considered (Hill, 1985);

comparing: the cost-effectiveness of the alternatives - cost probably measured in terms of the call on limited resources, and effectiveness assessed in terms of the potential learning experience;

choosing: at least how to start, and probably with options for continuing in different ways based on progress as it is made;

doing: which can be any one, or combination of the five styles of evaluation already outlined as the five 'fingers'.

As in the substantive decision-making process, these activities occur in no particular sequence, and are therefore arranged most easily in cyclic form. Figuratively, somewhat in the same way, the locus of the activities can be identified. Thus the scanning and doing modes of work are located in the 'Field of Operations'. And as the field of operations is, in this context, the substantive decision-making process itself, they can be seen to lie on the main circle in the diagram at the interface of the two circles. The remaining four methodological activities can then be seen as forming the palm of the hand from which the five fingers spring (see Figure 3).

Further, it is likely that the same sort of pairing of activities will occur between the methodological choice activities, as does between those of the substantive choice process. However this could easily become an infinite regression, and a halt must be called somewhere. As it is not intended to explore the methodological choices in any great depth in this paper, it will be helpful to explore aspects of the processes involved only briefly.

In the methodological decision process, just as in the substantive decision process, evaluation plays a vital, all-pervading role. Similarly also, review of the sort of criteria which might be used captures the essence of that evaluation. Grouped together these can be seen fundamentally as cost-effectiveness. For example, the cost in terms of the required resources would probably be measured in:

- the number of people with appropriate skills;
- the amount and quality of relevant data required;
- the style and capabilities of techniques and software;
- the type and capacities of computer hardware.

In most cases, such resources are limited, and they can become constraints on the

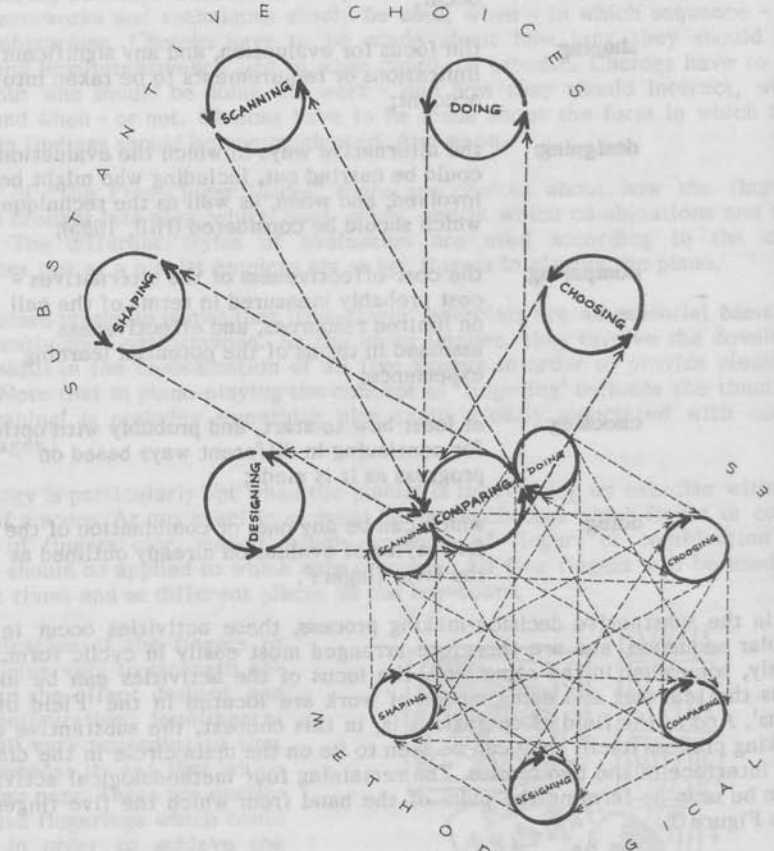


Figure 3. **The full "Five-Finger Model"**

range of choice open to the evaluator if there is not enough time and/or money to obtain them.

However, in certain circumstances, they can be improved, supported or even replaced for money:

- more effective people (training, consultants, etc);
- more relevant data (surveys, research, etc);
- bigger and faster technological support (computers, programs, etc);
- and so on.

However, this is always assuming that enough time - perhaps the ultimate constraint - is available.

The effectiveness criteria will be associated with the potential learning effect - the possible gain in understanding by those whose task is the decision-making and decision-taking. Such a potential learning effect can probably be assessed only indirectly using the accuracy, rigour, transparency, simplicity and speed of the alternative approaches as indicators.

END NOTE

Ideally, some sort of guidance should be provided as to how and when to use which style of evaluation. However, it seems difficult even to described some of them - let alone how and when they should be used. Optimistically, it can be hoped that "The Five-Finger Model", plus the idea of the 'five-finger' exercise, will be enough for a start. Perhaps attempts at using them in practice will provide the opportunity for better understanding - and out of that may grow the next statement offering guidance about the processes of evaluation.

Note

* The ideas in this paper first mooted in discussion with Nat Lichfield during the ECE colloquium on "Recent Developments in Planning Methodology", held in the Hague during October 1977. This was elaborated in "Evaluative Moments in a Planning Process" (the title of the paper originally submitted by the author to the international workshop in Delft, The Netherlands, February 1984), the second part of which was heavily revised just before and during its presentation.

REFERENCES

Bentham, J. (1967) **An Introduction to the Principles of Morals & Legislation** (originally 1789), Pergamon, Oxford (1967)

Blommestein, H.J., Bressers, J.Th.A. Bressers, A. Hoogerwerf (eds.) (1984) **Handboek Beleidsevaluatie: een multidisciplinaire benadering,** Samsom, Alphen aan den Rijn.

British Broadcasting Corporation (1985) **Radio Times,** 11th May 1985, London.

Van de Graaf, R (1985) Paper in this volume.

Hickling, A. (1974) **Managing Decisions - the Strategic Choice Approach,** Mantec, Rugby.

Hickling, A. (1982) Beyond a Linear Iterative Process, In: Talbot, R.J. (ed), **Changing Design**, Wiley, Chichester.

Hill, M. (1985) Paper in this volume.

Lichfield, N, P. Kettle and M. Whitbread (1975), **Evaluation in the Planning Process**, Pergamon, Oxford.

Lichfield, N. (1985) Paper in this volume.

Miller, D. (1985) Paper in this volume.

PART TWO: DUTCH APPLICATIONS

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

STRATEGIC CHOICE IN LPG POLICY

R. van de Graaf
Ministry of Economic Affairs
Directorate-General for the Industry

1. HISTORICAL OVERVIEW

Coming out of the seventies LPG emerged as a major potential source of energy and chemical feedstock. LPG stands for liquid petroleum gas. It consists of either propane or butane or a mixture of both. At room temperature under atmospheric pressure the substance is gaseous; it is transported either under pressure (approx. 17 atm.) or frozen (approx. -40°C).

Also in this decade public and political concern about the risk of industrial activities (among others: the Canvey Island study, the COVO study), became apparent, while national policy with respect to management of risks was still in the developing stage. In 1978, a calamity at Los Alfaques involving over 100 casualties underlined the risks involved in the transport of LPG.

Were we gambling on frozen fire ? Public and regulatory concern about the risk was quite understandable; the benefits of LPG were also well understood (typically in the LPG case, those persons enjoying the benefit are not the same as those at risk). Government at local, provincial and national level felt strongly involved.

The Ministry of Environmental Affairs chose the LPG-problem as a kickoff point for their policy on public safety, which gave LPG policy a more far reaching impact than it would have had if it had only concerned LPG. In this stress field LPG policy was constructed.

The first policy plan ("Aanlandingsnota") for LPG was put forward in 1979 and was based on the following assumptions:

* LPG is an attractive chemical feedstock, good for large scale use;
* there will be a surplus of LPG in the 1980's, ergo large volumes;
* pipeline transport is the safest option for transportation of LPG;
* large scale use can be made basically safe, small scale use is less safe.

Consequently, the intention of this policy plan was to:

* concentrate all LPG imports in one location in the Rijnmond area;
* limit all modes of LPG transport except pipelines;
* advocate the large scale use of LPG in petro-chemicals.

In Appendix I the policy plan is presented in a "commitment package".

What then became clear was the following. Firstly, the expected surplus of LPG in the 1980's turned out to be more like a shortage, therefore, the price was too high for LPG to be interesting as a large scale chemical feedstock. Secondly, the large volumes did not appear; the project in the Rijnmond area never got off the ground. Instead a terminal project got started up in a location logistically less suitable for pipeline transport. Thirdly, subsequent risk analysis demonstrated that pipeline

transport of LPG is not safer than other transport modes.

It could be said that the above mentioned LPG policy was succesful because in a formal sense the contingency plan had worked, and policy was changed as the situation changed. However, everyone involved had the perception that this policy was unsuccessful for at least three reasons. A potential large capital investment in the Rijnmond area did not come about. In addition, the essentially restrictive LPG policy had given some companies reason to invest outside the Netherlands. Finally, partly due to the above mentioned reasons, relationships between government and industry were bad, as were the relations between ministries (a losing team is bad tempered).

The situation with which we were faced when the option to use strategic choice arose, was the following:

* a complex problem with many interrelated decision areas;
* all decision areas in some way or other carried weight in a negotiation process;
* the negotiators distrusted each other and were extremely rigid in their contacts;
* most important: a common frame-work was lacking: each actor had his own perception of the problem, and of the uncertainty involved, and how that was to be solved.

2. INTERDEPARTMENTAL POLICY MAKING

In Holland, as presumably in other countries, ministries tend to become centres of power; they recognise some ministries as traditional allies, and others as traditional opponents. They are always considered to be alien actors.

Ministries meet each other when they make policy in a given field. They will claim (an amount of) responsibility for the field and they are quick to recognise the spin-off the subject field has into related policy fields.

Interdepartmental work sometimes resembles the game of diplomacy - or, for that matter, real life diplomacy. The subject under discussion has come to be the war theatre, the objective is to gain power, by control of policy or by precedent (spin-off into related policy fields).
In the LPG case the relation to other policy fields was quite strong. Where the basic question would seem to be: "What to do about LPG ?", for some it becomes:
- LPG is basically a dangerous good. Beginning with LPG, how do we make dangerous good transport policy ? (Ministry of Transport and Public Works)
- LPG is a public risk, and as such it is a precedent to set up policy on public safety (Ministry of Spatial Planning and Environmental Affairs; Environmental Department)
- LPG has spatial consequences, and as such it is a precedent to set up policy on zoning regulations for dangerous activities (same Ministry; Physical Planning Department)
- LPG is an economic commodity, and it may not be used as a precedent to introduce regulatory measures into a free market economy (Ministry of Economic Affairs).

I do not wish to dwell on this too long. However, I hope I have made clear that an interdepartmental team working on a policy problem, does not naturally operate in a teamwork mode, but in negotiation. The fieldunder discussion is wider than the field directly related to the technical problem connected with safe LPG operation.

In this sense you may sketch the situation. Each actor sees a larger field, but he must make his progress in the region where the fields overlap. The actor's interest is primarily his constituency in his own field, and he must relate the results of his negotiations back home: he is in the first place a negotiator.

The negotiators at the table will see their problem in this way: the table is ground and the private perception is figure as in Figure 1. This can be characterized as follows: the negotiators are working but they are not interacting; the subjective element is strong.

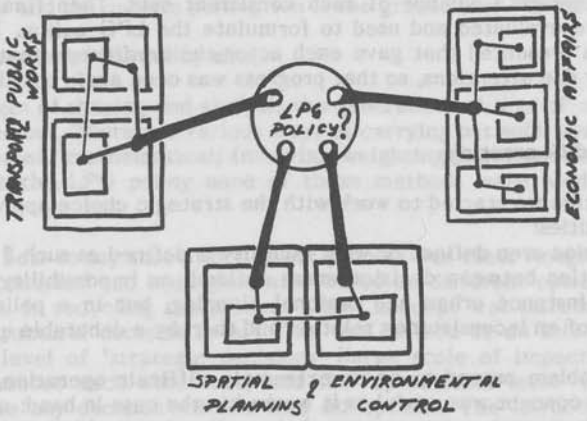


Figure 1. **The conventional organizational structure**

3. STRATEGIC CHOICE

3.1. Introduction of the approach

In the first two paragraphs I have briefly illustrated the complexity of the problem at hand, and how it is further complicated by its typical setting in the inter-departmental circuit. It was at the point in time when the above mentioned problems had brought progress to a stand-still that the Strategic Choice Approach was introduced to us. The first introduction was basically centered on the management of uncertainty which was indeed a point of discussion.

In the LPG case there were some rather basic uncertainties, but they were not recognised as such. There was the problem of risk: the amount of risk involved was an uncertain quantity, and even given the results of a risk analysis, the acceptability of the risks was an unknown quantity. Then there was the uncertainty with respect to the economic consequences of a restrictive risk management policy. These uncertainties were managed by making a scenario of the possible future and then believing that this scenario would model the real future. This kind of uncertainty management resulted in a debate on whose scenario was the most realistic, i.e. a confrontation between the 'certainties' of parties with conflicting interests. Interests in such a case are never common, what is needed is a common framework which gives each actor the confidence that his interest will get their due attention.

3.2. The cyclic progress

Originally the process was tackled in a linear sense: analysing one problem after the other, and trying to move from one decision to the next. In the nature of the LPG problem, decision areas were related and this approach did not work because a decision in one area had precedents in other decision areas; therefore no decisions could be taken.

A breakthrough came with the concept of the cyclic process. Make decisions, run through the whole problem, and come back again. Make and remake decisions until a consistent set is found. This 'analysis of interconnected decision areas' (AIDA) was used to generate a number of such consistent sets. Then finally these sets (scenarios) were evaluated and used to formulate the LPG policy. In this way a framework was presented that gave each actor the confidence that his interests would get their due attentions, so that progress was once again possible.

3.3. Shaping and Generating

In this perspective we started to work with the strategic choice approach. We soon ran into difficulties:
- how is a decision area defined or what quantity is defined as such ?
- how is a relation between decision areas defined: an impossibility relation may be real in for instance urban and regional planning, but in a policy decision it becomes more of an inconsistency relation and thereby a debatable quantity.

Shaping the problem proved to be an extremely difficult operation. Also in this case the cyclic concept was useful as it worked in the case in hand: a step forward in time was taken; each actor presented the policy he would chose if he had the liberty to go ahead by himself. This served to illustrate how close to compromise the actual situation had come and that a common goal could indeed be identified.

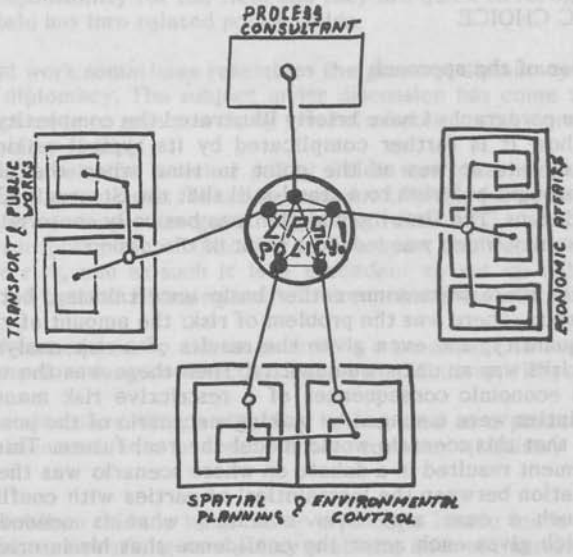


Figure 2. **Organizational structure from a strategic choice viewpoint**

What had happened ? I believe that the use of a common framework, together with a view on the problem as a cyclic process underlined the idea that the problem could be solved by analysis and the generation of solutions, rather than by the win or lose game of negotiation. This gave the actors a feeling that they were working together as members of a team. The team would generate solutions and would later present these to the various constituencies. Typically in this period there was a marked improvement in the social interaction among the actors.

The negotiators at the table have become members of a project team. The table is now the figure and the private perception is the ground (see Figure 2).

3.4. The evaluation of the policy sets

After the stages of shaping and analysis the generation of sets of consistent decisions is carried out. There are various ways of carrying out such evaluations, some of them extremely mathematical, involving weighting factors, matrix operations, and such. For the LPG policy none of these methods were used in an explicit sense.

It is probably best to say that a sorting out process was used. Roughly one hundred policy sets (scenarios) had been generated based on different options of strategic decision areas. In recycling the process a new concept was introduced, which we called the 'structural decision level'. This was defined as an intermediate level, between the level of 'strategic decisions' (large scale of impact), where every decision influences each other one, and the 'operational level' (small scale of impact), where any decision will fit into all options. The 'structural level' is a level which interacts with one strategic decision area.

In fact of course we were reshaping decision areas (clustering) by which a large reduction in the amount of scenarios was reached. In a following cycle we realised that effectively there were two main areas in the problem which were relatively weakly connected. By sectioning the problem into separate fields a final reduction down to eight scenarios was reached. In a learning process throughout successive cycles the problem was restructured.

The eight scenarios were all discussed. Largely due to the selection of contrastng scenarios, we had 'safety oriented' scenarios and 'economy oriented' scenarios and 'some in the middle'. The choice was then almost naturally made for the scenarios containing compromise, the 'ones in the middle'. The resulting policy is outlined in appendix II.

What in fact had occurred was that the process of evaluation or, preferably, choice hadbeen made implicitly. In the linear process, evaluation is the final step. In the cyclic process the shaping, the generation of alternatives, and the ranking exercises take place a number of times, and the consequences of making choices also influence the shape, etc.; in fact evaluation is present all the time; it is a part of the learning process.

4. FINAL REMARKS

The value of the strategic choice approach as we used it is not really given by the fact that it is a fine theoretical tool for generating and evaluating all possible solution sets to a complex problem, but much more by the common framework and

the cyclic process that come with it and their effect as tools in the interaction between interested parties working on a problem. These tools are effective in stimulating a process in which the energy of the actors is focussed on a solution oriented framework.

I believe the approach presented here can be questioned, and mainly on two points:

* if a rational approach should be directed at identifying the 'best solution', then is not this practical approach of strategic choice in danger of missing the best solution ?
* is the cyclic treatment of evaluation a good tool or is it a manner of manipulating the problem shape until it proves that the chosen solution is the best ?

I believe debate on such questions, even though it may seem academic, is valuable because these questions spring from a perception of subjectivity versus objectivity: a 'best solution' is an 'objective quantity' and a practical, or chosen solution is a 'subjective' one. While we often wish to present our solutions as objectively best (and unquestionable), I believe that the solution that contains the subjective choice of a wide group of actors is the real goal and as such the closest approximation of what we call objectivity.

APPENDIX I

An excerpt of the commitment package for Dutch national LPG policy (1979)

Action set

Actions:
Concentrate LPG imports by:
- Give license to a major LPG terminal in Rijnmond area;
- Restrict terminals in other areas;
- Advocate large scale use of LPG in petro-chemical plants.

Explorations:
- Investigate possibilities and problems in LPG pipeline transport.

Policy set

Delayed action:
- Start legislation on LPG transport in order to be able to limit all modes of transport except pipeline;
- Start a large scale risk analysis into all activities with LPG;
- Initiate a government committee to incorporate the results of the risk analysis into integral LPG policy.

Contingency:
- Initiate a working committee on interim policy to handle contingencies as they arise.

APPENDIX II

An excerpt of the commitment package for Dutch national LPG policy (1984)

Action set

Actions:
- traffic regulation: increase regulatory measures for LPG ships on the main national waterways;
- LPG inland tankships: adapt design and construction in order to effectively prohibit large LPG outflows;
- LPG railcars: large contingents should always be shipped in block trains; in case (for smaller contingests) of shunting some extra measures are specified;
- LPG road tankers: extra design adaptations are introduced to further decrease probabilities of LPG outflows;
- Stationary installations: specifications for zoning around such installations are given;
- Road transport routes: will basically be limited to the Dutch dangerous goods network. New stationary installations will be located in places compatible to this network.

Explorations:
- research into the sage domains around seaships will be done;
- the collision safety of seagoing tankers carrying pressurised LPG will be analysed;
- the possibilities of repressing fire around LPG inland tankers will be analysed;
- further analysis with respect to effective and economic means of prohibiting Bleve's on LPG road tankers will be conducted.

Policy set

Delayed actions:
The result of the above mentioned analyses will be implemented in the relevant policy decisions and it will be introduced in national regulation and in international regulatory bodies.

Contingency plans:
- in case an LPG-terminal in Amsterdam is conceived then a refuge location for LPG tankers will be constructed;
- in case international regulatory bodies do not accommodate the proposed measures then the general character of such measures must be guaranteed along other lines;
- in case the analysis with respect to probiting Bleve in road tankers does not give satisfactory results then other means will be used to create an acceptable situation;
- in case measures with respect to LPG filling stations prove to be intractable due to local situations, then a decision will be made regarding other necessary measures or removal of the filling station.

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

SOME EXPERIENCES WITH PROJECT APPRAISAL IN THE DUTCH CENTRAL GOVERNMENT

A. Sorber
Ministry of Finance, The Hague
Dept. of Policy Analysis

1. INTRODUCTION

This paper deals with the relevance of ex-ante evaluation (project-appraisal) as one of the more important instruments of policy analysis for public decision-making, concentrating on experiences with ex-ante evaluation (procedures) in the Dutch central government. As background information, section 2 gives a short overview of the most important institutional aspects. The next section sketches the basic approaches to project appraisal as a methodological scheme of reference. Section 4 forms the core of the paper. It starts by outlining the main elements of an appraisal study. With the help of this framework, recent experiences with respect to methodological and procedural aspects of ex-ante evaluation are discussed. The final section contains some conclusions and suggestions for making policy analysis more effective.

2. INSTITUTIONAL ASPECTS

The official introduction of policy analysis in The Netherlands in the seventies was to a large extent the result of the early enthusiasm about the use of the Planning Programming and Budgeting System during those years in both the United States and other countries (Gray and Jenkins, 1983). In 1971, the Minister of Finance established an interdepartmental commission to guide and stimulate policy analysis in the Dutch central government. The Commission, composed of top level civil servants from various ministries and chaired by the Director General of the Budget, developed a number of methods and techniques, particularly with respect to ex ante evaluation. Consequently, guidelines on cost-benefit analysis were published (Norma 1, 1974; Norma 2, 1975), training programmes and seminars were organised and, under the auspices of the Commission and with the support of its Secretariate, the ministries performed a substantial number of studies.

By the end of the seventies, however, it became clear that the philosophy of a centralised approach turned out to be no longer successful. On the basis of an extensive evaluation of its functioning, it was decided in 1982 to abolish the Commission and to create a new unit for policy analysis as part of the directorate of Budgetary Affairs of the Ministry of Finance. At the same time, the centralised approach was abolished and, in accordance with the new philosophy, the ministries themselves were considered to be fully responsible for their policy analyses. The role of the newly created department of policy analysis focused on infrastructural aspects. Along these lines the training facilities were enlarged, the development of methods and techniques concentrated on operational aspects, and facilities were created for consultancy activities. The application of policy analysis was further stimulated by publication of guidelines and technical reports (Financiën, 1983; Financiën, 1984) and the publishing of the quarterly "Beleidsanalyse" ("Policy Analysis").

A complete new element was the initiative to set up a national platform for the

exchange of views and experiences with policy analysis between both central and local government, as well as the academic world and the private sector. The rationale behind the platform-concept is that the interplay between theory and practical experience is of essential importance for the (further) development of realistic approaches in the field of policy analysis. To promote the idea of the platform, the Minister of Finance invited a group of decision-makers and analysts involved in policy analysis, representing more or less the four sectors mentioned above, for a conference on the state of the art of policy analysis. As a follow up, a core group has started organising a program of workshops, seminars etc..

3. METHODOLOGICAL ASPECTS

In order to start the discussion of the methodological aspects of project appraisal it may be worthwhile to look at the various techniques that are available. At the outset it should be clear that the significance of evaluation studies is "restricted" to giving the decision-maker (more) insight in the problems at hand and the consequences of the analysed alternative solutions. The final decision is up to the policy-maker, the methods and techniques are not suited to give the ultimate answers.

Generally speaking, project appraisal methods can be categorized into two groups: monetary methods and non-monetary methods. The main difference between these two lies in the possibility or impossibility of expressing the effects in monetary terms. The non-monetary methods can be subdivided into simple effect-surveys and multicriteria techniques. The last category takes into account the "weights" of the various impacts of the project.

With respect to the monetary methods, the most important subcategories are cost-benefit analysis and cost-effectiveness analysis. The first one focuses on contrasting the costs of the project with the benefits and the third party effects. Characteristically, both costs and effects are expressed as much as possible in monetary terms. Cost-effectiveness analysis concentrates only on one side of the "balance sheet". In most cases benefits are taken for granted and only the costs of the various alternatives are analysed.

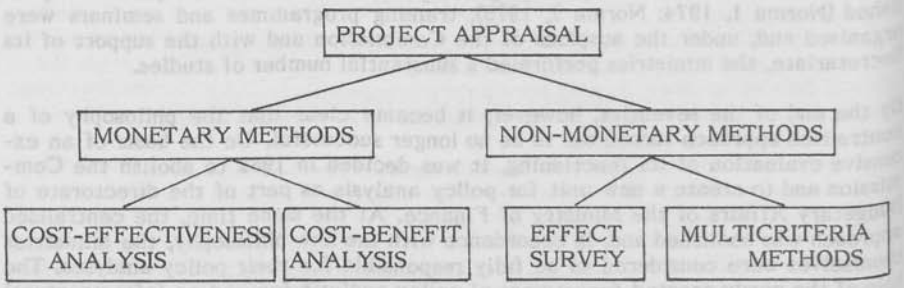


Figure 1. **Ex ante Evaluation Methods**

The monetary approach was considered to be the most promising up until the mid-seventies. Well-known examples in The Netherlands are the cost-benefit analyses of a new railway-line to Schiphol-airport (Schipholllijn, 1970) and of a second na-

tional airport (Central Planningbureau, 1974). A much discussed example of a cost-benefit study in quite another sector is the analysis of employment projects for the handicapped (Haveman, 1977). More recently cost-benefit analyses have been performed on land reclamation projects by means of the Markerwaard-polder (Markerwaard, 1983).

Regarding cost-effectiveness analysis, the study of recreation facilities in the Grevelingenbekken (1976) and the more recent study of a wind energy project (Lievense, 1981) are interesting examples.

A very striking example of an effect-survey approach is the score-card method developed by the Rand Corporation and used in a study on the protection of the Oosterschelde estuary from floods (Rand, 1977). As to the multicriteria methods, one finds many examples in the field of urban and regional planning (Help, 1983). Very recently, an interesting study on the setting of priorities for highways was published (Rijkswaterstaat, 1984). The interesting point in this case was the willingness showed by the Minister of Transport and Public Works in providing the 'weights' that are indispensable for such a multicriteria analysis.

4. EXPERIENCES

4.1. General procedure

For the discussion of experiences with project appraisal in the central government, we will take as a scheme of reference the general procedure for performing ex ante evaluation on a public project. This procedure can be indicated by a sequence of steps. The scheme does not imply, however, that in any actual analysis one always starts with the first step and necessarily goes straight on to the next one. In practise (as we will see in the next subsection), one often starts "somewhere in the middle" and has to return one or more times to previous steps as the study proceeds. The general procedure for ex-ante evaluation is indicated in Figure 2 (Financiën, 1983).

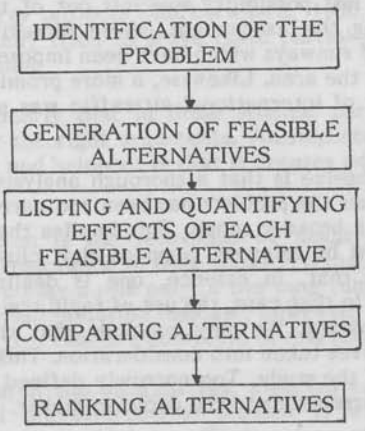


Figure 2. **General Procedure for Ex-Ante Evaluation**

This sequence of steps gives a starting point for our discussions on the practise of project appraisal.

4.2. Experiences with appraisal of public projects

In many cases, assignments for ex-ante evaluation are cast in terms of a cost-benefit analysis. This does not imply, however, that the decision-maker is fully aware of the specific characteristics of this method in comparison to other relevant appraisal methods. This point requires thorough consultations between the analyst and the decision-maker. It should be made clear to the decision-maker that in cases where most of the effects cannot be expressed in money terms, it will be inappropriate to use a monetary approach. Furthermore, in cases where the decision-maker is not prepared to indicate the weights of the various effects, it is questionable if a multicriteria method will be adequate. Sometimes a modified method may be appropriate. This was the case in a study by the Ministry of the Interior on a nation-wide uniform telephonenumber to alarm police, fire brigade and hospital facilities. This approach appeared to be very adequete.

Another important aspect of the assigned task has to do with the given range of alternatives that are to be analysed. In the extreme situation, only one more-or-less obvious solution to the policy problem has to be studied. In cases where the analyst accepts his assignment without question, he is, in fact, ignoring two fundamental steps of the whole evaluation procedure: the identification of the problem and the generation of feasible alternatives. This can have very awkward consequences. At the end of the study, a new discussion might begin about completely new alternatives that are considered to be more relevant by the policy-maker. There should be consultations between the analyst and the policy-maker on this specific topic to avoid unnecessary delays.

The cost-benefit analysis of the second national airport is, to some extent, an example of a study where the task was rather narrowly defined (Central Planningbureau, 1974). The (expected) growth of air traffic in The Netherlands could be absorbed by a second national airport or by extending the existing Amsterdam Airport 'Schiphol'. The last possibility was left out of the study because of the directives accompanying the assignment (these directives originated from the limitations on the use of runways which have been imposed on Schiphol as a result of the noise pollution in the area. Likewise, a more prominent role for the regional airports in the handling of international airtraffic was not taken into consideration.

A related point to emphasize is that a thorough analysis of the indicated policy problem may indicate that the problem has been too narrowly defined whereas, in fact, the problem is of a broader scope. This implies that a completely different set of alternatives might be relevant. Suppose that close examination of an air traffic problem reveals that, in essence, one is dealing with a more general transportation problem. In that case, the use of rapid trains may be as appropriate as enlarging airport capacity. We may conclude that the scope of the problem determines the alternatives taken into consideration. This, in turn, is very decisive for the final outcome of the study. Too narrowly defined alternatives will lead to solutions that might be irrelevant to the decision-maker.

Regarding the identification of the effects, we will first look at the cost-item. As funds become more and more limited these days, this aspect receives more attention. As to the identification of the costs of a project, attention should be paid to the 'full costs principle' (Quade, 1975). This implies that both investment costs as

well as maintenance costs, during the lifetime of the project, should be taken into account. Omitting maintenance costs, although they may be relatively low on a yearly base, can lead to important miscalculations because of the relatively low increase in labour productivity in that sector (Openbare Uitgaven, 1984). On the other hand, 'sunk costs' can be left out of the picture (Quade, 1975). One has to be very sure, however, that the means of production concerned cannot be used in any alternative way. In addition to the costs, the benefits have to be identified. In a social cost-benefit study, the effects on third parties must also be taken into account.

As stated before, the choice of method in a specific case depends on the scope of the actual task. It should be clear, however, that the outcome of the study also depends to a large extent on the specific assignment. The Schiphol railway line, mentioned before, was analysed both with a business-like cost-benefit approach and with a social cost-benefit analysis. Although the first study led to negative outcomes for the railway company, the last one showed a positive outcome (mainly due to net benefits for passengers). Although for decisions of the central government, insight into the operating (net) revenues of a project is very useful (e.g. in the case of an airport, one may think of landing and take-off dues), information on the social net benefits (positive and negative effects on third parties) may sometimes be of greater importance.

To get the 'complete' picture of all the effects is often a very complicated matter. In the first place, thorough expertise is required with respect to the subject. There exists, however, no unique answer to the question of how many and which effects have to be taken into account in a specific case. Moreover, double counting has to be prevented, and one should be very cautious to reflect redistribution effects, which are sometimes more important for the decisionmaker than the efficiency effects (Leman and Nelson, 1981), in the right way. In the second place, there is the problem that the various participants are focused on getting onto the balance sheet the specific items that are favorable for the desired outcome, which may be quite different for the various parties concerned, and on omitting negative aspects. During the (earlier) Markerwaard study, the advocates of the project emphasized again and again (additional) items as housing, an airport, and wind energy projects (Markerwaard, 1980, 1983). As for the opponents, they systematically tried to prove (with partial success) that there was not any need for these items.

This example indicates clearly that in these studies, participants are forced to search very seriously for the right arguments. An important consequence of this built-in system of checks and balances is that it creates additional possibilities for a higher quality analysis.

A special case to mention is the relevance of negative third party effects. Generally, there is some reluctance to take these effects into account, especially when they are expected to play a role only in the long run. The problems we have nowadays with water and soil pollution indicate, however, the far reaching importance of performing analyses that are complete in this respect.

For the initial comparison of the alternatives, a survey can be set up which contains an overview of all the alternatives studied as well as their effects. In general, it is not necessary that all these effects are quantified.

With respect to the ranking of the alternatives, a decision has to be taken about the appropriate appraisal method. As explained earlier, the approach to be chosen depends on the possibilities of translating the effects into monetary terms and/or

the willingness of the policy maker to indicate the weights that are necessary for a multicriteria analysis. Finally, there is the possibility of a modified analysis that concentrates only on certain costs and benefits.

For a further discussion of the ranking of the alternatives, we will now concentrate on cost-benefit analysis. As stated earlier, the outcome for a particular alternative is not restricted to concepts as 'net benefit' or 'benefit-cost ratio' but also includes those aspects that cannot be translated in monetary terms - e.g.'increase in social-psychological well-being' (Haveman, 1977). This implies that, in many cases, a ranking in a technical way is not possible and it is up to the politician to 'weigh' the alternatives.

An often heard reproach of decision-makers has to do with the fact that ex-ante evaluation has only a limited importance, as things may work out in a way not foreseen by the analyst. Although the last part of this statement is certainly correct, it is unreasonable to held the analyst responsible for not perfectly forseeing the future. On the other hand, this view makes it necessary to pay due attention to the aspect of uncertainty. A rather simple and effective way to cope with this problem is to perform a so-called sensitivity-analysis. This implies that, for the basis hypotheses underlying both the initial description of the policy problem and the calculations, different values are used to construct alternative scenario's for the costs and benefits of the project. In the study of the second national airport, this procedure was followed with respect to the anticipated growth of airtravel (Central Planningbureau, 1974). This approach may also be used in cases where questions arise on the magnitude of the social discount rate, future prices and shadow prices for unemployed resources. In the studies of the Markerwaardpolder this procedure was widely used. If it is relevant to know the outcomes in very bad circumstances, it may be worthwile to calculate the socalled 'worst case' on the basis of unfavorable values for the most important variables.

Finally, it should be stated that in many cases, studies take much more time than expected. This is often due to a lack of good project management of the study. Apart from unnecessary and costly delay, one may end up with outcomes that have become irrelevant for the decision-maker (Goemans en Smits, 1984).

5. CONCLUSIONS

For an ex-ante evaluation to be effective, there has to be an absolute clarity about the scope and character of the policy problem at hand and of the potential alternative solutions. This implies elaborate consultations between the decision-maker and the analyst about the assigned task right from the start of the study.

At the same time, decisions have to be taken about the performance of the analysis. The time available and the availability of research facilities are decisive factors for the actual study. In many cases, there is no room for a performance according to scientific standards. Sometimes a 'quick and dirty' approach will suffice (Leman and Nelson, 1981; Verdier, 1984).

Special attention has to be paid to the management of the analysis. One of the most striking points is that, right from the start, all parties concerned should be able to participate in the study. In this way, expensive delays can be avoided. In an elaborate study there is often a conglomerate of steering and working groups. An effective project management may be of great importance for finishing the study on time and for staying within the limits of the research budget (Goemans and Smits, 1984).

It goes without saying that the presentation of the analysis always needs special attention (Financiën, 1984).

Finally, analysts should always be aware of the limitations of any policy analysis. It is up to the politicians to make the final decisions. Evaluation studies are only aimed at providing structured information and relevant insights.

REFERENCES

Central Planningbureau (1974) **Cost-benefit analysis second national airport in the Netherlands: General Report**, Monograph nr. 18, Central Planning Bureau, The Hague.

Financiën (1983) **Beleidsonderzoek; het ontwikkelen en beoordelen van beleidsmaatregelen en -projecten**, Rapport van de werkgroep Projectanalyse, Ministerie van Financiën, Staatsuitgeverij, The Hague

Financiën (1984) **Evaluatiemethoden, een introductie**, Rapport van de Afdeling Beleidsanalyse, Ministerie van Financiën, Staatsuitgeverij, The Hague

Goemans, T. and H.N.J. Smits (1984) Kostenbeheersing van een mega-project: de Oosterscheldewerken, **Economisch Statistische Berichten** 69, nr. 3478 (24 october), 988-996

Gray,A. andB. Jenkings (1983) **Policy Analysis and Evaluation in British Government**, Royal Institute of Public Administration, London

Grevelingenbekken (1976) Kosteneffectiviteits analyse van de inrichting van het Grevelingenbekken, Interdepartementale werkgroep kostenbaten analyse Grevelingenbekken, Lelystad.

Haveman, R.H. (1977) **A Benefit-Cost and Policy Analysis of The Netherlands' Social Employment Program**, University of Leiden, Dept. of Economics

Help (1983) **De Helpmethode voor de evaluatie van landinrichtingsprojecten; beschrijving en verantwoording**, Landinrichtingsdienst, Staatsuitgeverij, The Hague

Leman, K. and R.H. Nelson (1981) Ten Commandments for Policy Economists, **Journal of Policy Analysis and Management**, Vol. 1, nr. 1, 97

Lievense (1981) **Windenergie en waterkracht**, Rapport van de Begeleidingscommissie Voorstudie Plan Lievense, Staatsuitgeverij, The Hague

Markerwaard (1980) Directeuren-generaal van de Rijkswaterstaat, de Rijksplanologische Dienst en de Rijksbegroting, Beleidsanalyse Markerwaard, The Hague.

Markerwaard (1983) Economische analyse van de ontwikkeling van het Markerwaardgebied: samenvatting van een achtergrondstudie opgesteld door de interdepartementale stuurgroep Markerwaardeconomie, Ministerie van Verkeer en Waterstaat, Lelystad

Norma 1 (1974) The First Interim Report on Norms and Standards 2 for Cost-Benefit Analysis, **Beleidsanalyse** 74-1 (available in English Translation)

Norma 2 (1975) Measurement and Valuation - Second Interim Report of the Working Group on Norms and Standards for Cost-Benefit Analysis, **Beleidsanalyse** 75-4 (available in English Translation)

Quade, E.S. (1975) **Analysis for Public Decisions**, American Elsevier, New York

Rand (1977) Protecting an Estuary from Floods - A Policy Analysis of the Oosterschelde (prepared for the Netherlands Rijkswaterstaat)/ B.F. Goeller (et al.), Vol. I-VI, The Rand Corporation, Santa Monica

Openbare Uitgaven (1984) Besluitvorming grote waterstaatkundige projecten in Nederland, **Openbare Uitgaven** 16, August 4, pp. 165

Rijkswaterstaat (1984) Het planningsproces van rijkswegen, Rapport van het Ministerie van Verkeer en Waterstaat (Hoofddirectie van de Waterstaat), The Hague

Verdier, J.M. (1984) Advising Congressional Decision-Makers: Guidelines for Economists, **Journal of Policy Analysis and Management**, Vol. 3, nr. 3, 421

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

MULTICRITERIA EVALUATION AS A PLANNING INSTRUMENT: REMARKS FROM THE VIEWPOINT OF NATIONAL PHYSICAL PLANNING

J. v. Staalduine
National Physical Planning Agency, The Hague *

1. INTRODUCTION

Decision-making in physical planning nearly always involves the weighing of various interests. Per interest, or if one prefers, per criterion, the value of alternatives can be determined, or at least placed in an order of preference. Preferences, however, are not comparable for all criteria; consequently physical planning is an ideal field for multicriteria evaluation, and in recent years, at a national level, several attemps to apply this method in decision-making have been made.

In this paper I will discuss some of these attempts. First, the interorganizational aspect of physical planning is explained, since this is important when assessing the applicability of evaluation methods. Next, three cases are briefly treated, two of which took place in the National Physical Planning Agency, and one in consultation between various departments.

The practical experience gained can be differentiated into experiences gained in internal planning on the one hand, and in external decision-making on the other hand. The concluding section of the paper deals with possible perspectives for further development.

2. THE INTERORGANIZATIONAL ASPECT OF PHYSICAL PLANNING

Physical planning is the search for and development of the best conceivable adaptation of the physical environment and society in the interest of society. From this description, it is immediately clear that physical planning is a multi-facetted activity which is not limited to any one department or administrative level. The contributions of various departments and authorities are essential when dealing with spatial questions. This makes it necessary to coordinate the activities of various policy areas in such a way as to promote optimum spatial development.

The aim of physical planning is to clarify the relationship between policy decisions and the physical environment, weighing all aspects as well as possible so that the decisions taken will yield the best possible results for society and the environment. Spatially relevant decision-making is not confined to single authorities, but results from a process of negatiation and consultation between authorities. Decision-making is, moreover, distributed over various levels of government. This interorganizational aspect of physical planning has important consequences for decision-making procedures and for the design of policy preparation processes.

In the Netherlands, an effort has been made to structure the coordinating activities which are of importance to physical planning by designating areas of

* The views expressed in this article only represent the author's opinions

government policy as either facets or sectors, as well as by designing rules for horizontal coordination (within one administrative level) and vertical coordination (between various administrative levels). Since this structuring is important to the design of the policy preparation process, we will devote some attention to it.

This resulted in a Report, called the Report of the 'De Wolff Committee' (1970), which exerted considerable influence on policy formulation in the area of physical planning. The Report distinguishes five, wide areas of government policy, namely:

- international relations
- furtherance and distribution of prosperity
- spatial planning
- science and technology
- social and cultural wellbeing.

These are the so-called facets. Attention is focused on one particular aspect of various activities, and every effort is made to achieve the integration of all government activities from this single point of view. On the other hand, sectors, such as traffic and transport, drinking water supply etc., are concerned with the complete programming of one branch of government activity (which corresponds with a department or branch of a department).

The differentiation into sectors and facets has proved to be a useful instrument for denoting the place of policy memoranda in the whole of government activities, and for defining responsibilities. It should, however, be realized that the choice of what is considered a facet or a sector may change over time. It is not only determined by the range of the field of interest, but is also dependent on social circumstances. The division into facets and sectors depends on the viewpoints one considers most important for arranging government policy. An example of this is energy supply, which undoubtedly used to be considered a sectoral activity, but now has the attributes of a facet.

There is no hierarchical relationship between facet and sector organizations. This results in the socalled 'two-track' idea showing the relation between sectors and the spatial facet. Decision-making with respect to spatially relevant matters proceeds along two lines: the facet line, and the sector line. Each has its own special characteristics and adheres to its own rules. The sector cannot produce complete spatial assessments, since its area of atention covers only part of the total, physical, planning field. Neither can the sector be absorbed into the facet. Sectoral interests are of prime importance in the sector line. In relation to the sectoral interests, in many cases the spatial aspect plays only a modest role. This does not imply a lack of objectivity, but is a consequence of responsibilities. A sector cannot be absorbed in the physical planning facet, since within this sector other, nonspatial factors also operate. Naturally, the simultaneous existence of two lines of policy formation may lead to conflict. In order to avoid or to solve these conflicts, it is essential that it be clear at which level, how, and by whom the lines of facet and sector should be tied together.

There must be both horizontal and vertical coordination. Horizontal coordination means harmonizing the spatially relevant plans within one administrative level, while vertical coordination means the harmonization between different administrative levels. At a national level, the 'structural outline sketches' and 'structural outline plans' are important. The former are facet memoranda which relate to the general, spatial developments which the government desires to guide in a certain direction. In view of the wide scope of the material to be dealt with, the government decided not to publish the Third Memorandum on Physical

Planning in one volume, but in the form of a series of reports. At the moment, the Third Memorandum consists of: The Orientation Report, the Urbanization Report, and the Structural Outline Sketch for Urbanization, as well as the Report on Rural Areas and the Structural Outline Sketch for Rural Areas. In addition to these reports and structural outline sketches, there are also structural outline plans. These can be described as reports containing the main lines and principles of a certain sector of government policy. They are of general importance for national spatial policy, and yield spacial insight into the spatial aspects of the essential facilities in that sector. These facilities also include concrete projects and natural, culturalhistorical, and scenic values which should be preserved. Structural outline plans, consequently, constitute the intersection between the sector and the spatial facet.

Naturally reports, outline sketches, and outline plans are not separate entities. The reports and structural outline sketches create the framework within which the structural outline plans should be developed. The policy which is presented in a structural outline plan will then have to fit in with, or be an elaboration of, the general spatial views presented in the structural outline sketch. The sector Minister concerned has the primary responsibility for the structural outline plan. The Minister of Housing, Physical Planning, and Environment is jointly responsible, as the coordinating Minister for Physical Planning. It is his responsibility to ensure that the spatial consequences of sectoral policy are in harmony with spatial facet policy, and he is also responsible for the effects of the projects on physical planning.

It is against this background that decision-making with respect to physical planning takes place. Any evaluation methods used must therefore meet the requirement of being applicable to policy preparation processes in which various authorities participate. A distinction can be made between internal and external application.

The above-mentioned criterion is naturally, of particular importance for external application, but is also relevant in the case of internal application. For example in the National Physical Planning Agency, a multitude of internal interests operate, interests which must be reflected in the chosen evaluation method.

We will deal with two instances of internal application, followed by an example of an external application of multicriteria evaluation.

3. INTERNAL APPLICATION OF MULTICRITERIA EVALUATION

In recent years, the National Physical Planning Agency has acquired experence in applying evaluation methods to help choose between alternative routes for highways. On several occasions, it has made use of multicriteria evaluation to help in developing its position on a particular route. In order to appreciate the relevance of this practical experience, it must be realized that the evaluation was carried out on the basis of material received form the Ministry of Public Works (the report on road routes), as well as other avialable material.
In the evaluation of road routes, two alternative methods were constantly applied simultaneously (weighted summation technique and concordance analysis). These methods are extensively described by Hordijk **et al.** (1977) and Voogd (1983).

3.1. Principles of the methods

Briefly, in the weighted summation technique the scores per criterion, per alternative, are standardized and multiplied by the weight of the criterion.

The products are then added up, per alternative. The largest sum indicates the preferred alternative, the next greatest sum the second most preferred alternative, etc. By means of sensitivity analysis, it can be determined which combination of changes in weights brings about a change in the sequence of preference between two alternatives. If this 'degree of resistence' is large, the sequence of preference is fairly stable since it might be altered in the case of even relatively small changes in weights. Moreover, a sensitivity analysis can also be carried out per criterion. In that case, the question is to establish which change in value of one particular criterion brings about a change in preference sequence.

In concordance analysis, pairs of alternatives are compared, and the relative superiority of one alternative over another is determined on the basis of the value of the weights, and the degree of difference in criterion scores. One alternative is considered superior to another if the criterion upon which that alternative scores higher, is considered important. If, for instance, five criteria are used, and alternative A scores higher than alternative B with respect to two criteria, but the weights of these two criteria are very high by comparison with those of the three criteria for which B scores higher, then it is assumed that alternative A is superior to alternative B.

The same applies with respect to differences in scores If one alternative scores marginally higher with respect to a certain criterion than another, this will carry little weight in the determination of superiority; if the difference is large, it does carry weight. After determining the preference order per pair, it is simple to determine a general preference sequence. The alternative which has been found superior the greatest number of times in the comparison of pairs is placed at the top of the list, the alternative which was found superior slightly less often is placed second, etc.

In concordance analysis, insight into sensitivity is also essential. The degree of resistence is a number which indicates the difference in dominance between tow alternatives. This number runs from 0.0 (no difference) to 1.0 (maximum difference, that is to say, for each criterion, the dominant alternative scores better than the non-dominant alternative). Here too, it is possible to perform a sensitivity analyses per criterion, namely for the comparison between two alternatives.

See the earlier-mentioned article by Hordijk **et al.** for a formal description of these methods.

3.2. The results of the application

The preceding methods have been applied in determining the road route RW 2 Den Bosch Eindhoven. It would be going too far to discuss the whole question here in detail. It will suffice to say that the existing road (RW 264), which runs through Vught outside of Boxtel and through Best, can no longer handle the traffic. A new

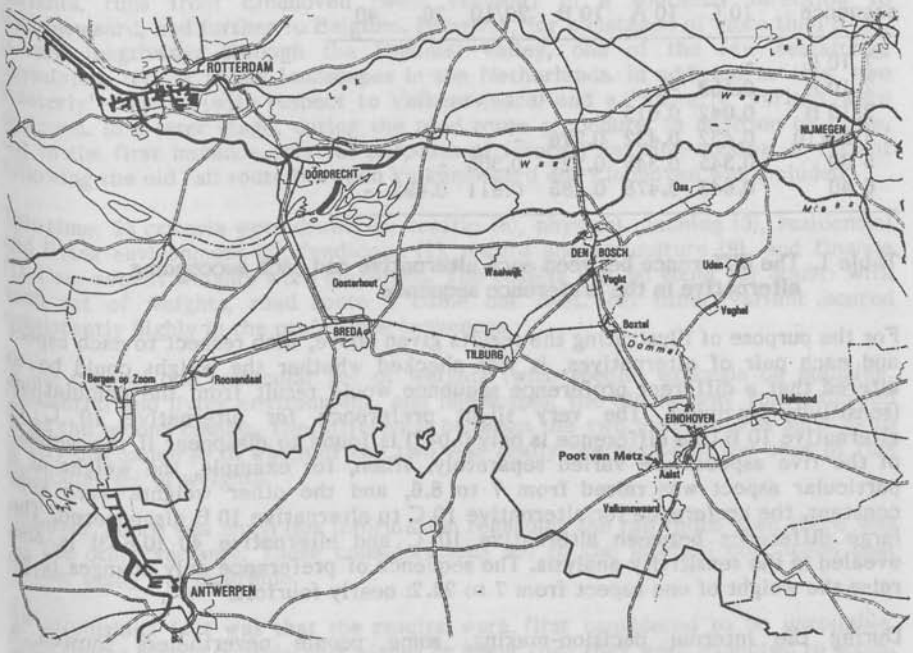


Figure 1. **A map of the study area**

route was proposed, but had to be reconsidered due to objections on the bases of the expected environmental effects. As a result, six additional variants were considered in the road route report:

- variant 10 A: consisting of RW 264 (enlarged to a motorway), with an easterly bypass near Best
- variant 10 B: in which the whole of RW 264 is enlarged into a motorway
- variant 10 C: ditto, but with a westerly bypass near Best
- variant 20: which is the old road route for RW 2 and involves considerable environmental objections
- variant 20/10: which is a combination of the variants 20 and 10
- variant 40: finally, which is a more easterly route outside of Eindhoven.

Eighteen criteria were dwawn up, divided into traffic aspects (7 criteria), physical planning aspects (4), residential and living environment (2), rural environment (2), agricultural (2), and financial (1). These criteria were given one weight (intersubjectively determined) and scored per variant (ordinally). The results supported the opinions in favour of the various variants 10, which were already

being developed. Between the variants 10, there was little difference to be found, but the differences were large with respect to the other variants. To illustrate this, table 1 shows the resistence and the sensitivity analyses of the concordance analyses, as reported at the time.

Preference sequence	10 C	10 A	10 B	20/10	20	40
1 10 C	-					
2 10 A	0.153	-				
3 10 B	0.042	0.015	-			
4 20/10	0.432	0.432	0.276	-		
5 20	0.345	0.345	0.227	0.305	-	
6 40	0.643	0.478	0.485	0.911	0.492	-

Table 1. **The difference between each alternative and each succeeding alternative in the preference sequence**

For the purpose of illustrating the results given above, with respect to each aspect and each pair of alternatives, it was checked whether the weight could be so altered that a different preference sequence would result from the calculations (sensitivity analysis). The very slight preference for alternative 10 C to alternative 10 B (the difference is only 0.042) is found to disappear if the weights of the five aspects are varied separately. When, for example, the weight of a particular aspect was raised from 7 to 8.6, and the other weights were kept constant, the preference for alternative 10 C to alternative 10 B disappeared. The large difference between alternative 10 C and alternative 40 (0.643) is also evealed in the sensitivity analysis. The sequence of preference only changes if we raise the weight of one aspect from 7 to 25.2: nearly fourfold!

During the internal decision-making, some people nevertheless showed a preference for variant 40, which multicriteria analysis has shown to be less favourable. The results of the multicriteria analysis were evidently not enough to limit the discussion to a choice between 10 A, 10 B and 10 C. Instead, on the basis of the table of criteria it was discussed whether the score per criterion, and weight per criterion had been correctly assessed. The arguments for and against the various routes were weighed against the evaluation matrix. In the process, it was found that, in addition to the criteria mentioned, advocates of variant 40 took into account an additional criterion, namely the connection with the (inter)national roads network.

The other variants scored lower than variant 40 for this criterion, which had not been included in the multicriteria analysis. The question was whether the multicriteria analysis should be carried out anew, this time including the lastmentioned criterion.

In the second round of internal decision-making, it was decided, however, not to use this additional criterion (or rather: it was not permitted to weigh the balance). The ultimate choice was in favour of the variants 10.

It is worth mentioning that the external decision-maker (the Minister of Public Works) ultimately chose the variant 10 B, after being advised by the Consultation Committee for Roads and the Pulic Works Council.
We will expand further on our example in NBrabant, viz. the Extended Poot van

Metz. The existing motorway 269/69 runs through the centres of Valkenswaard and Aalst. This is a single carriageway road, intended for fast and slow traffic, while in the builtup areas, cycle paths adjoin it. Partly for this reason, a new road was considered necessary. For this road, a route memorandum with several variants was drawn up.

Road route 1 of the socalled extended Poot van Metz, one of the proposed variants, runs from Eindhoven (west tangent) in a southerly direction to Valkenswaard, and further to Belgium. However, for a distance of more than 8 km, it runs lengthways through the Dommel valley, one of the few remaining, unviolated, rivulet valley landscapes in the Netherlands. In addition to this, two 'easterly' variants (with respect to Valkenswaard) and a 'westerly' variant were designed. In a later stage, during the road route procedure, in addition to these, but in the first instance only for purposes of internal decision-making, a variant following the old rail route between Valkenswaard and Eindhoven was included.

This time, 24 criteria were drawn up: traffic (6), physical planning (3), residential and living environment (5), landscape (1), nature and agriculture (9), and finance (1). Four sets of weights were employed. The results left no room for doubt: with each set of weights, road route 1 came out best. No other variant scored consistently highly in the preference sequences.

In the case of one particular set of weights (with special exphasis on planning, residential and living environment, and landscape), the old railway route showed up as the second choice. The difference between the two variants was such as to justify discussion as to which of the two was really preferable in the case of this particular set of weights.

However, there was much less information regarding the rail route than there was for the other variants. These considerations led to the internal conclusion that further study was desirable.

An astonishing fact was that the results were first considered to be unreliable, since routes scored considerably higher or lower than they were expected to score. Further study of the evluation matrix, however, showed these results to be correct, thereby increasing the insight into the problem. Though the users had themselves drawn up the matrix, they had not been aware of certain relative advantages and disadvantages of various routes.

The road route has not yet been decided upon. The Consultation Committee for Roads is still dealing with the matter.

4. EXTERNAL APPLICATION OF EVALUATION METHODS

The multicriteria technique 'Electra' was applied in drawing up the Structural Outline Plan for Military Sites. It was used in order to facilitate the choice between 25 possible locations for 6 company practive sites.

The working method was as follows: fifteen criteria were drawn up, as well as eighteen sets of weights. The weights could only assume the values 0, 1, and 2, or sometimes 1.33 and 2.66. As a consequence of the large number of sets of weights, and the limited variation in values which the weights could assume, the evaluation technique was applied to a spectrum of matrices including nearly all information sets except those which were clearly absurd.

The results were in keeping with the approach: eighteen sometime completely different preference sequences resulted, from which, in the decision-making process, ultimately six locations were selected. The method used was not to select a certain set of weights, and on the basis of this the most preferred locations, but rather to choose locations which regularly had high scores. The technique thus acquired a purely indicative significance. In addition to this, in the choice of locations, one objective was to achieve a certain geographical distribution. The choice was conditioned by this fact, and, since this conditioning element had not constituted part of the technique, it was necessary to adapt the results to this.

Naturally, geographical variants can also be invisaged. In the ultimate decisionmaking for the Structural Outline Plan for Military Sites, various 'packets' of company practice sites were drawn up, the socalled models, which were then weighted against each other. No carefully outlined technique was employed, but rather alternatives and criteria were considered and weighed.

The locations included in these packets were, in principle, the locations which had resulted from the Electra analysis, supplemented with a few locations which were later found to be of importance and had not been included in the evaluation method.

5. CONCLUSIONS

A clear distinction should be made between evaluation matrices (information regarding a plan, structured around alternatives and criteria), and an evaluation technique (placing alternatives in an order of preference on the basis of information from an evaluation matrix). Both of these are found to be important in internal plan formation. The drawing up of an evaluation matrix is, naturally, always relevant when alternatives must be designed and compared, but the use of an evaluation technique may also improve the quality of planning. It was found that, in the case of the Extended Poot van Metz, the application of the technique led to greater insight into the evaluation matrix, especially concerning the relative pros and cons of the alternatives. This is not surprising.

In practice, collected information is often checked by subjecting it to a slightly different working method. A column of figures is added up from top to bottom, and from bottom to top. If this yields a difference, the calculation must be carried out once more. Comparing results has the effect of improving quality. It should, therefore, not be surprising that the application of an evaluation technique to an evaluation matrix can, in many cases, provide deeper insight. The matrix can be altered on the basis of this new insight, and this, in turn, offers more possibilities for ultimate decision-making. This is not all. In at least one practical instance, the practice sites, it proved possible to employ an evaluation technique as a sort of language, as a result of which the choice between locations could be submitted to discussion. That, ultimately, results are interpreted as one sees fit, handly detracts from the fact that by employing this technique, it was possible for planners from various departments jointly to compare and weigh alternatives in a systematic manner.
The situation is different in the case of decision-making. Here, the evaluation technique really does not play a role. Sensitivity analysis results are not taken into consideration. In practical situations, this was not done even when it would have been convenient for the decisionmaker. Also the resistence measure did not function in decision-making, and the sequence of preference was mainly used as an opening for discussion, a point of departure.

What was discussed in detail in the decision-making, was the content of the evaluation matrix. This is to be expected. A road route, or a practice site, involves certain disadvantages which may render decisionmakers open to criticism. In such cases, they must be highly conscious of the advantages, in any case for the purpose of making their own judgement, in order to arrive at a well-founded decision. The road route which, at a first glance, appears suitable, but is poorly assessed by the evaluation technique, must be well documented.

In the practical cases, it was striking that the weights ascribed to criteria beforehand, even in the case of alternative sets of weights, often proved to be completely incorrect when it actually came to decision-making. When faced with reality and the necessity to make a choice between, for instance, better landscape and less costs, the weights change per project, dependent on the degree of 'better' and 'less' of the area in question, on earlier decisions, and on political pressure.

In order to arrive at sound decision-making, it is, therefore, very important to be in possession of an efficacious evaluation matrix. During the policy preparation which takes place interorganizationally in physical planning, this can be qualitatively improved by using an evaluationtechnique. This should be so well designed as to be usable by persons who approach policy preparation from various angles. The technique must be simple in structure, contain few or no normative assumptions, and the results should be interpretable on the basis of the information from the evaluation matrix.

One of the consequences of the condition that the results of evaluation techniques must remain interpretable, is the strict desirability for evaluation not to be applied to complex plans, but for the plans to be first thoroughly unravelled (hierarchy and interdependency of decisions), and for evaluation techniques to be applied only to certain parts of the plan. Application to a complex plan in its totality is apt to yield preference sequences or indices which cannot be traced back to basic information, and consequently can make no further contribution to the improvement of the quality of the plan.

REFERENCES

De Wolff Committee, Rapport van de Commissie Voorbereiding Onderzoek Toekomstige Maatschappijstructuur, Tweede Kamer, Zitting 1970-1971, nr. 10914, Staatsuitgeverij, The Hague (1970)

Hordijk, P., Staalduine, J. van, and Voogd, H., Twee methoden van planevaluatie, **Planning, methodiek en toepassing**, 2 (1977) 11-24

Voogd, H., **Multicriteria Evaluation for Urban and Regional Planning**, Pion, London (1983)

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EVALUATION METHODS IN REGIONAL PLAN MAKING: THE CASE OF THE REGIONAL PLAN FOR TWENTE

J.W. de Vos
Physical Planning Agency, Province of Overijssel, Zwolle
Currently Economic-Technological Agency, North Holland

"The new methods, properly used, release everyone from the tyranny of imposed ideas" (J.C. Jones)

1. INTRODUCTION

An important aspect of urban and regional planning is the weighing of one interest against another. Since space, in a given area, is limited, not all claims can be completely satisfied. New housing, for example, is built at the expense of agriculture and can also have negative consequences for the natural scenery. In turn, natural scenery can have a major influence on the proposed route of a motorway, etc. In short, a choice must be made: either one function or another, or something in between. To clarify this, and to rationalize the process of choosing, it has become customary to formulate a number of alternatives. These alternatives can be seen as examples of the way in which conflicting goals and interests can be weighed.

Even with the use of alternatives, it is in no way certain that the different goals and conflicting interests will be equally represented during the discussions. This will depend both on the method of generating the alternatives, as well as the selection of a limited number of them to represent all the possible alternatives, i.e. the entire range of policy options. In both steps of the planning procedure, implicit preferences may influence the process. A planner may often begin the planning process with a whole range of preconceived, though unconscious ideas about the area under consideration. Opinions about which areas are suitable for housing, or which areas are of science value, may well prejudice his decisions.

In developing urbanization models for regional planning, such a problem exists in a more general way. In many cases, planning is still concerned with searching for new locations for housing or industry, or for railways and roads. For the most part, this demand for interests of agriculture, nature, and landscape have taken back seat to the above interests. With regard to urbanization, the rural area is always on the defensive. It could be that this will never change. Although most urbanization studies do take into account the 'green' interests, 'red' has the upper hand. The principles of urbanization supply the a priori criteria for the design of alternatives; the consequences for the rural area provide the a posteriori criteria for the evaluation of the alternatives. What is to be evaluated is determined by the criteria for design. Because of this, alternatives which are advantageous for the natural scenery and for agriculture can easily be overlooked.

A cyclic planning process can mitigate this difficulty to some extent; however, for a number of reasons, it appears nearly impossible to make more than one cycle in practice, let alone more than one. In this way, primacy is given to urban space; while the rural area, as the supplier of space, nearly serves to 'balance the books'.

A method is outlined below which deals with the conflicting interests of 'red' and 'green' in a balanced way (see (1)). The method consists of four simultaneaous, multicriteria analyses for the following categories of land use: housing areas, industrial and office areas (subsequently to be called working areas), agriculture,

and scenery, combined with an optimization procedure. The method can be extended to other forms of land use. For the sake of simplicity, we will confine ourselves to the four main kinds of land use mentioned above.

With the help of Figure 1, the method will be illustrated using the urbanization study for the 1984 revision of the regional plan for Twente, dating from 1966 (see: (2)). The region of Twente, located in the eastern part of the Netherlands, measures approx. 560 sq. mi. In 1983, it had a total population of 560,000 (60% living in the central urban zone), and a working population of 175,000 (65% in the central urban zone). Along with the eastern part of Groningen and the southern part of Limburg (provinces in the Northern and Southern Netherlands respectivily), Twente is the third region with severe socio-economic problems, mainly due to industrial decline. The region contains large areas of great scenic beauty.

2. THE SUITABILITY OF AN AREA FOR DIFFERENT FUNCTIONS

It is clear that not all locations within a certain area are equally suitable for fulfilling one of the four functions: housing, working, agriculture, or scenery. One location may have a greater number of qualities for a given function than another. As convercely, it will depend on the proposed function in determining which requirements have to be met by a location. Therefore it must be established:

- which factors, or so-called criteria, determine the suitability of a location in fulfilling a given function:
- to what extent a location satisfies these criteria?

With these problems, the following questions are relevant:

- do all criteria have the same importance for a given function, and if not,in what way is the importance distributed over the criteria;
- how is the planning area divided into locations;
- how are the locations assessed for the various criteria?

With the help of Figure 1, these and similar questions appertaining to the properties of the supply of space will be discussed.

3. CRITERIA

Which criteria are relevant in assessing an area for different types of land use, i.e. different spatial functions? Figure 2 shows the criteria that were used in the urbanization study for the regional plan for Twente. The criteria were obtained from a study of the literature, as well as professional judgement. Further down, we will return to the meaning of the digits and brackets used in Figure 2.

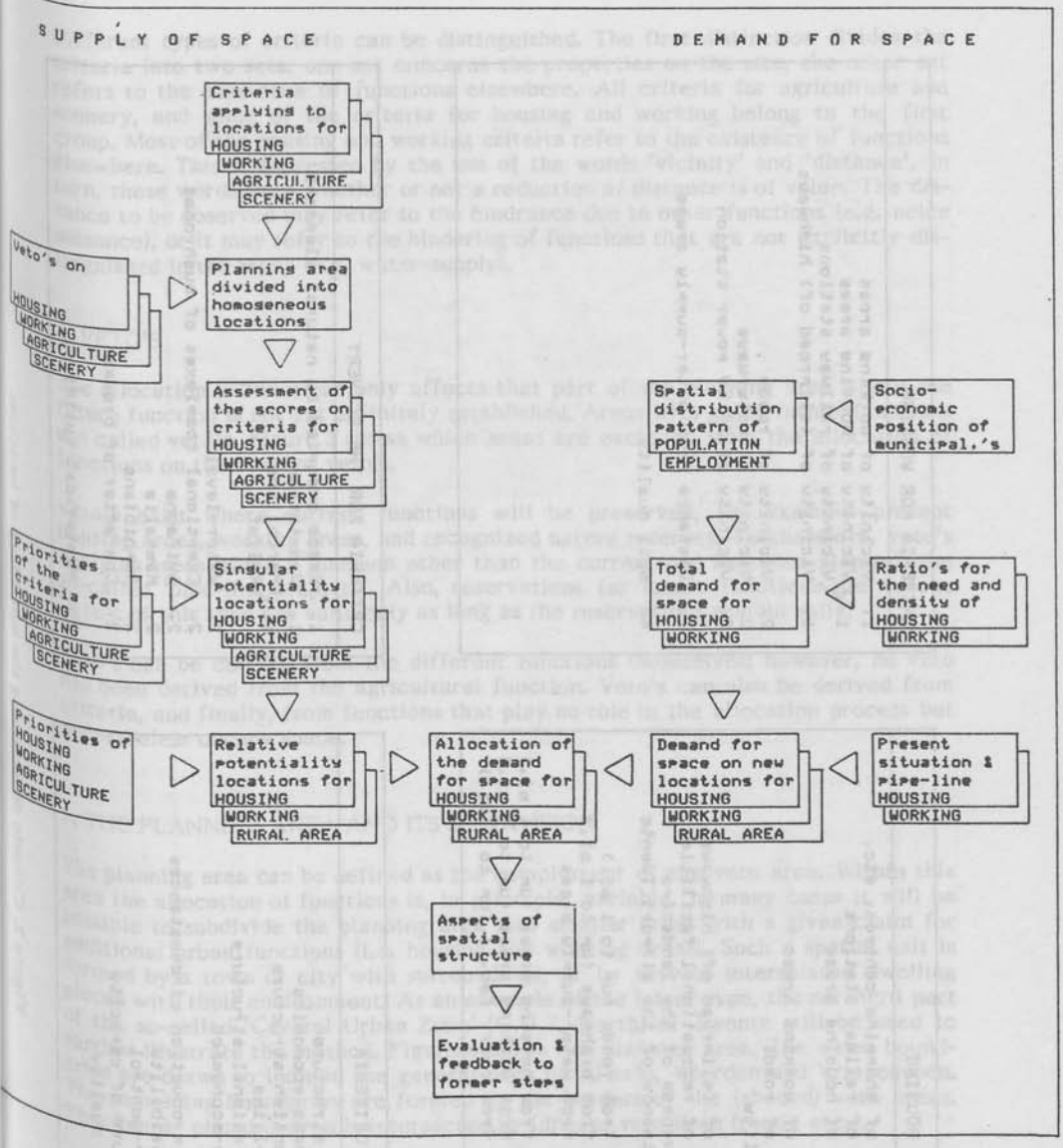


Figure 1. **The design of alternative urbanization models**

CRITERIA FOR HOUSING

1. Vicinity of shopping centres, etc.
1. Vicinity of railway stations
1. Vicinity of housing areas

2. Vicinity of working areas
2. Vicinity of roads
2. Soil quality

3. Vicinity of (sliproad of) highways
3. Vicinity of recreational facilities
3. Attractiveness of the landscape
3. Noise nuisance of the airport Twente

4. Distance from roads (up to 1 mi.)
4. Distance from highways (up to 1 mi.)
4. Distance from water-supply areas

5. Distance from housing areas (up to 1 mi.)
5. Distance from working areas (up to 1 mi.)
5. Distance from recr. fac.'s (up to 1 mi.)

CRITERIA FOR WORKING

1. Vicinity of working areas
1. Vicinity of housing areas
1. Vicinity of railway stations
1. Vicinity of (sliproad of) highways

2. Vicinity of roads
2. Vicinity of waterways
2. Vicinity of 10 kV power station

3. Distance from water-supply areas

4. Soil quality

AGRICULTURAL CRITERIA

1. Business structure
   - business economics
   - production structure
   - parcelling
   - occupational structure
   - socio-economics

2. External production circumstances
   - accessibility
   - water-control
   - measure of capacity
   - soil quality

CRITERIA FOR THE SCENERY

1. Space occupied by natural elements
1. Vegetation
1. Parcelling
1. Relief

2. Water-level
2. Exeptional complexes of mushrooms
2. Avifauna
2. Mammalia
2. Amphibians
2. Character of brooks

3. Historical buildings

Different types of criteria can be distinguished. The first distinction divides the criteria into two sets: one set concerns the properties on the site, the other set refers to the existence of functions elsewhere. All criteria for agriculture and scenery, and some of the criteria for housing and working belong to the first group. Most of the housing and working criteria refer to the existence of functions elsewhere. This is expressed by the use of the words 'vicinity' and 'distance'. In turn, these words show whether or not a reduction of distance is of value. The distance to be observed may refer to the hindrance due to other functions (e.g. noice nuisance), or it may refer to the hindering of functions that are not explicitly distinguished in the study (e.g. water-supply).

4. VETO'S

The allocation of functions only affects that part of the planning area where the future function is not yet definitely established. Areas with given future functions are called veto's. Figure 3 shows which areas are excluded from the allocation of functions on the basis of veto's.

Veto's exist where current functions will be preserved, for example present housing areas, working areas, and recognized nature reserves. Furthermore, veto's apply to areas where a function other than the current one has been planned; the so-called 'pipe-line projects'. Also, reservations for future functions are veto's. Veto's of this type are valid only as long as the reservations remain valid.

Veto's can be derived from the different functions themselves; however, no veto has been derived from the agricultural function. Veto's can also be derived from criteria, and finally, from functions that play no role in the allocation process but nevertheless occupy space.

5. THE PLANNING AREA AND ITS SUBDIVISION

The planning area can be defined as the complement of the veto area. Within this area the allocation of functions is, in principle, variable. In many cases it will be possible to subdivide the planning area into smaller areas with a given claim for additional urban functions (i.e. housing and working areas). Such a spatial unit is formed by a town or city with surroundings, or by several interrelated dwelling places with their environment. As an example of the latter type, the northern part of the so-called 'Central Urban Zone' (C.U.Z.-North) of Twente will be used to further illustrate the method. Figure 4 shows this planning area. The outer boundaries are drawn to include the general area of Almelo, Wierden and Vriezenveen. The remaining boundaries are formed by the borders of the (shaded) veto areas. This 'blank' planning area is intersected by (drawn) veto lines (roads, etc.).

The circumference of the planning area found in this manner, and its subdivision, usually result in zones too large to be considered as homogeneous. Homogeniety is needed to make possible an assessment of the zones using the different criteria. An ecological survey was used for this further subdivision. It provided a way of dividing the area which was suitable for assessment on the basis of the criteria for scenery, because it described the areas in terms of local characteristics.

V E T O'S	OPERATIVE FOR THE FUNCTIONS			
	HOUSING	WORKING	AGRIC.	SCENERY
A. Derived from the functions themselves				
* Veto areas				
1. Housing area		*	*	*
2. Working area	*		*	*
3. Woods/wild grounds	*	*	*	
B. Derived from the criteria				
* Veto areas				
1. Daytime recreational facility	*	*	*	*
2. Water-supply area	*	*		
3. Area of noise nuisance > 40K-units	*			
* Veto lines				
4. Road	*	*	*	*
5. Railway	*	*	*	*
6. Waterway	*	*	*	*
C. Derived from other functions				
* Veto areas				
1. Recreational facility (sojournment)	*	*	*	*
2. Refuse processing	*	*	*	*
3. Sand-pit	*	*	*	*
4. Storage basin	*	*	*	*
* Veto lines				
5. Power line	*	*	*	*

Figure 3. Areas excluded from allocation of functions

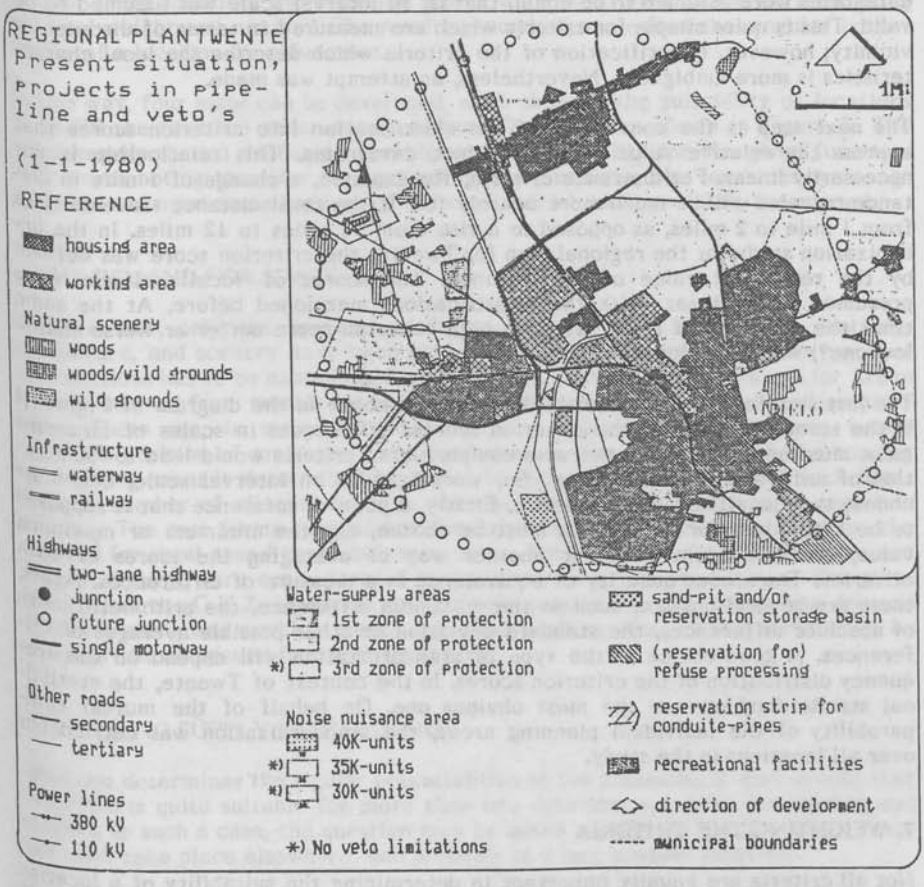


Figure 4. **The planning area of the Central Urban Zone-North**

In some cases, even such a subdivision must be refined. This may be necessary for relatively large or oblong zones. Such zones are hardly homogeneous with respect to their proximity to other zones. Since most criteria for housing and working functions are based on distance, a final subdivision was carried out. The smallest spatial units of the planning area are called 'locations'. Figure 5 gives the subdivision of the C.U.Z.-North.

6. ASSESSMENT OF THE LOCATIONS, OR DETERMINING CRITERION SCORES

The locations must now be assessed using all the criteria which measure their suitability for urban (housing and working) and rural (agriculture and natural scenery) functions.

In the context of the regional plan for Twente, all locations were classified into a

number of categories for each criterion. The differences between successive categories were assumed to be equal, that is, an interval scale was assumed to be valid. This is quite simple for criteria which are measured in terms of distance or vicinity; however, quantification of the criteria which describe the local characteristics is more ambiguous. Nevertheless, an attempt was made.

The next step is the conversion of the classification into criterion scores that express the relative value of the distinct categories. This relationship is not necessarily linear. For disstance criteria, for example, a change of 1 mile in distance traveled will be much more acutely felt if the total distance traveled rises from 1 mile to 2 miles, as opposed to a rise from 11 miles to 12 miles. In the urbanization study for the regional plan for Twente, the criterion score was defined by the reciprocal value of the distance. The scores of 'local' criteria were presumed to be linear, with the classifications mentioned before. At the same time, the direction of the score (is a high criterion score better or worse than a low one?) was taken into account.

The last step in this stage (the third block from above in the diagram of Figure 1) is the standardization of the criterion scores. Differences in scales of measurement must be eliminated, otherwise comparison of criteria would lead to 'summation of mills and woorden shoes'. So, working with an interval scale, one must choose two quantities of equivalence. Firstly, a point of reference that is supposed to be a standard for all criteria must be chosen, e.g. the minimum or maximum value, the arithmetic mean, or another way of averaging the scores of each criterion. The second quantity of equivalence is a measure of differences. Again, there are several options, such as the maximum difference, the arithmetic mean of absolute differences, the standard deviation or other possible averages of differences. A good choice of the type of standardization will depend on the frequency distribution of the criterion scores. In the context of Twente, the statistical standardization was the most obvious one. On behalf of the mutual comparability of the individual planning areas, the standardization was carried out over all locations in the study.

7. WEIGHTING THE CRITERIA

Not all criteria are equally important in determining the suitability of a location as a housing or working site, or the desirability to save its present agricultural function and its natural scenery. Some criteria are more important for a certain function than others. Figure 2 shows the order of importance of the criteria, using integers. Highly correlated criteria were combined into one criterion, which is reflected by the brackets. It is worth noting that the agricultural criteria were selected and weighed by an institute for agricultural economics (the L.E.I. at The Hague) in a way that was not made public because of the confidential data involved. The ranking of criteria for other functions is based upon a consensus of the planners. Fortunately,sensitivity analysis made it clear that the results were rather robust.

For the next step in the method (cf. Figure 1), a large number of arithmetical techniques are available. Since we were dealing with ordinal weights, it seemed obvious to use a qualitative multicriteria evaluation technique. However, techniques of this kind are rather complicated and require much computer time. Therefore, the simple technique of weighed summation was preferred. It therefore follows that the ordinal ranking of the weights must be converted into a series of cardinal numbers. This conversion took place, somewhat arbitrarily, by averaging all imaginable conversions that maintain the given rankings (i.e. by computing the

median point set defined by the ranking of preferences). For each of the functions, and for every location, the criterion scores were multiplied by the corresponding cardinal weight and summed. The resulting number was called the singular potentiality of a location for a function.

In this way, four maps can be developed, each showing the suitability of locations to fulfill the function under consideration. The maps are shown in Figure 6. The more heavily shaded a location, the higher its potentiality for the given function. It is worth noting that, for the agricultural potentiality, the locations were aggregated to larger areas because of the confidential data previously mentioned.

8. THE DEMAND FOR SPACE

Up until now, only the properties of the planning area for housing, working, agriculture, and scenery have been explored. Now, the space needed for each of the functions has to be examined. Figure 1 shows that the total demand for urban areas is derived from models for the spatial distribution of population and employment. These are, in turn, derived from the objectives regarding the socio-economic position of municipalities. Furthermore, the total demand for housing and working areas is determined by ratios for the need and density of both functions. A number of allocations have already taken place, the socalled pipe-line projects. The rest forms the demand for urban space at new sites. Finally, the 'demand for rural areas' is the balancing item. The rural functions, agriculture and natural scenery, were not separated because they overlap one another. The planning area of the C.U.Z.-North totals (in round numbers) 14,000 acres, the demand for new housing and working areas, 500 acres each, so that 13,000 acres must be found for the compound function 'rural area'.

9. MATCHING SUPPLY WITH DEMAND

When one determines the sigular potentialities of the locations, it may appear that a location is quite suitable for more than one function, e.g. both for housing and working. In such a case, the question may be asked whether it matters which function must take place elsewhere, and probably in a less suitable location.

One can imagine that value is attached to the fact that a given function is awarded the best spots, even if it is at the expense of other functions. Thus, it is possible, as in the case of the criteria, that different preferences are assigned to the four functions. These preferences may also be converted into caridnal numbers. Multiplications of the sigular potentiality with this number results in the so-called relative potentiality for each of the four functions. In this way, the potentialities of a location for the various functions are interrelated.

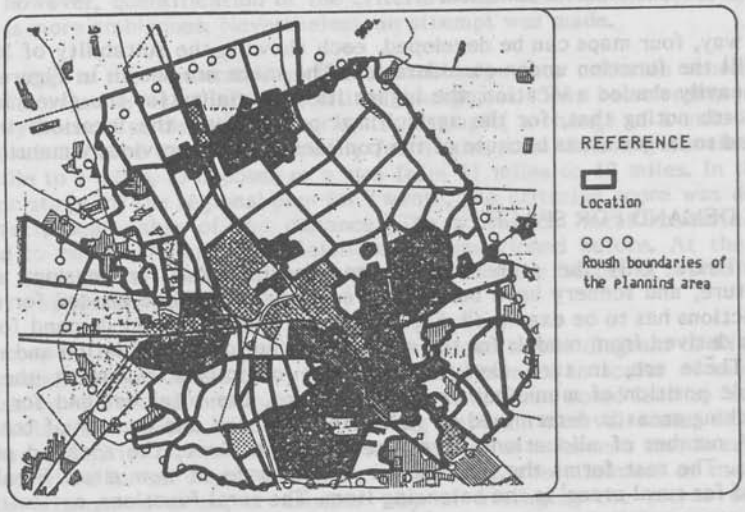


Figure 5. **Subdivision of the planning area into locations**

Before the ultimate matching of supply with demand for space, the two rural functions are linked because of their spatial overlap. Allocation of new urban functions means that present values of agriculture, or values of natural scenery, or both (on account of the overlap) will be sacrificed. In this connection, the relative potentiality of the compound rural function is defined as the maximum of the relative potentialities of both rural functions separately.

Now we have, on the one hand, a planning area subdivided into locations analyzed for their potential for housing, working, and conservation of present rural functions. On the other hand, we know the demand for housing, working, and conservation of present rural functions. In this phase of the method, the equivalence in handling conflicting interests is most clearly shown (i.e. the dotted rectangle in Figure 1). Not only the urban functions, but all functions, determine the distribution of functions over the locations. This implies that none of the functions is, a priori, favoured, and none is, a priori, the balancing item in the allocation process.

Confrontation of supply with demand in such a way that the maximum profit is yielded resolves itself into the well known operations research 'transportation problem'. Implementation for the C.U.Z.-North with four equal function weights leads to the result shown in Figure 7a. In order to get more insight into the second and third best solutions, allocation also took place for twice and thrice the demand for housing and working and the resulting space for rural functions.

As an example, Figure 7b gives an alternative allocation with emphasis on the housing function, while the other weights were kept equal. Comparison with Figure 7a, clearly shows the influence of emphasizing the housing function. All locations with the highest (singular) potentialities for housing (cf. Figure 6a) do, in

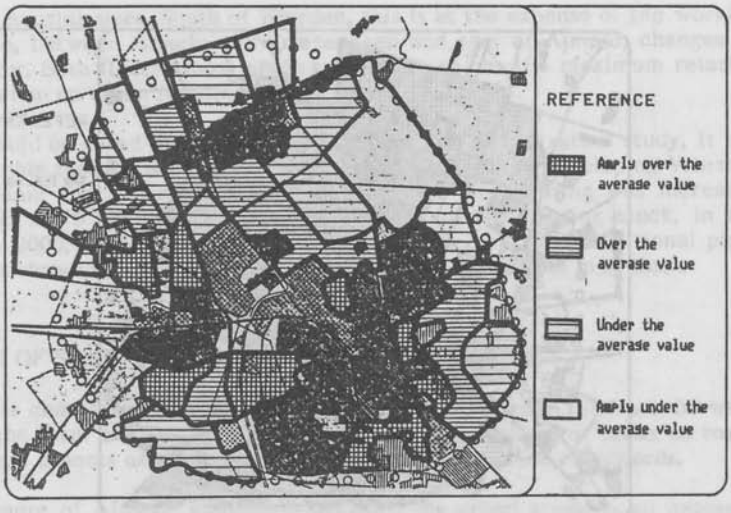


Figure 6a. **Housing potentialities in C.U.Z.-North**

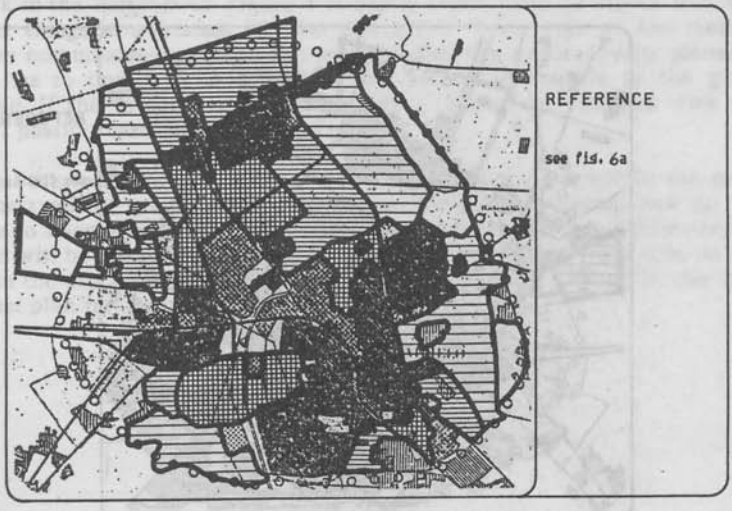


Figure 6b. **Working potentialities in C.U.Z.-North**

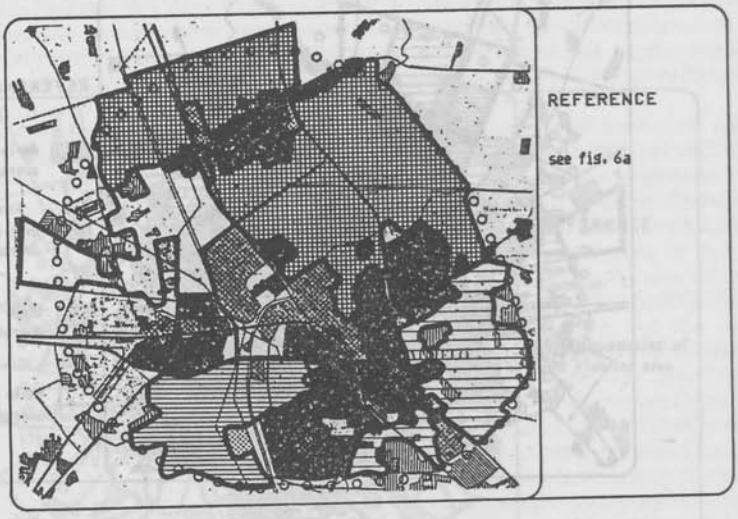


Figure 6c. **Agricultural potentialities in C.U.Z.-North**

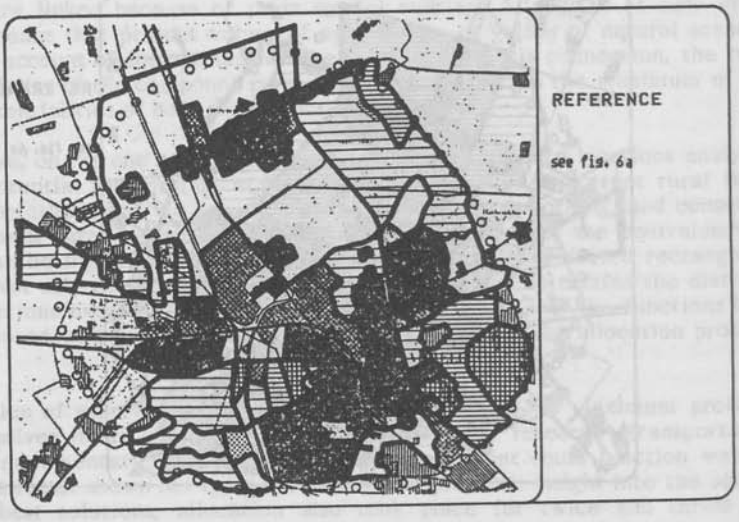


Figure 6d. **Potentialities of the scenery in C.U.Z.-North**

fact, become housing sites. South of Wierden, this is at the expense of the working function. Also, between Almelo and Vriezenveen and east of Almelo, changes of functions occur. Both figures show which rural function has the maximum relative potentiality when no urban function was allocated.

Finally, it should be noted that during the course of this urbanization study, it appeared inevitable that decisions would have to be made about developing Vriezenveen. This meant that since the study, the 'pipe-line' for housing was increased from about 65 to 75 per cent of the total growth of the housing stock, in the period 1982 - 2000, in this part of the Central Urban Zone. This additional pipeline project is shown in Figure 7 by the reference-mark 'as yet in pipe-line'.

10. ASPECTS OF SPATIAL STRUCTURE

Because of its character, the method is unable to incorporate relations between locations in the form of infrastructure, ecological relationships, or ideas on townplanning. These aspects of spatial structure must be introduced afterwards.

The development of Almelo, with lobes between the urban areas, is an example. Because of their potentiality for housing, the green areas of the town were urbanized in all of the allocations explored ('dwelling in the city park with shopping centres, etc. close by'; cf. Figure 6a). This urbanization was prohibited in later models by introducing a veto for housing. This veto is shown in Figure 7.

11. EVALUATION AND FEEDBACK TO PREVIOUS STEPS

The last block in the diagram of Figure 1 is really superfluous. It states that one must discover things in a process of trial and error. Never rely on the method itself, but use common sense. When the method does not produce very plausible results, one has to improve, i.e. simplify, the method if possible in the given period of time. If this is not possible, further results must be obtained with the help of other, possibly less sophisticated, methods.

In addition to other planning cycles using alternative assumptions within the same method, other types of questions have to be resolved. For instance, how do the results relate to known policies? Can the results be presented straightforwardly? Which models will be used for public discussion? In other words: what role do the results play in the process of policy-making? Some interesting findings in the case of the regional plan for Twente will now be discussed.

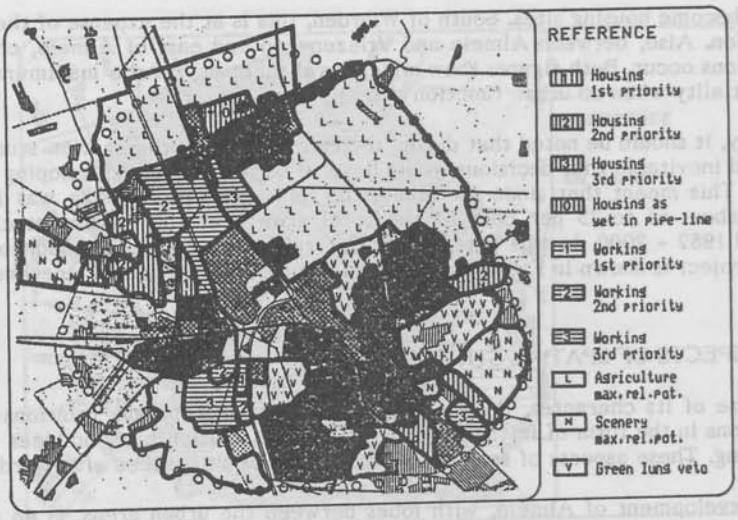


Figure 7a. **Allocation of functions with equal weights**

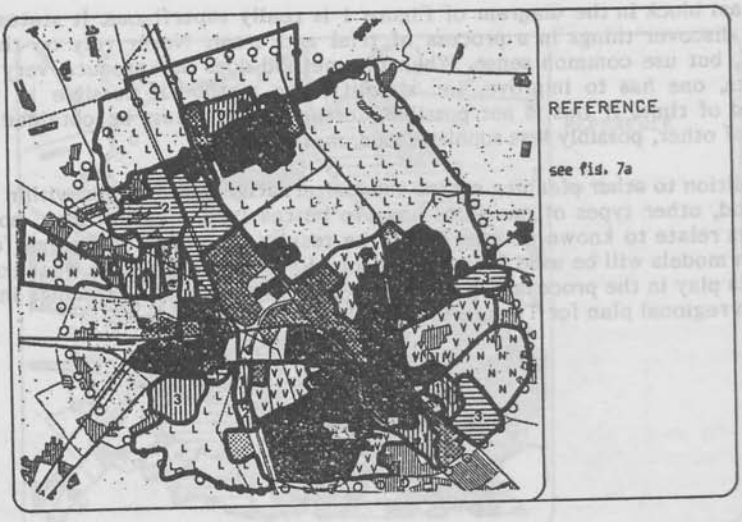


Figure 7b. **Allocation of functions with the emphasis on housing**

12. THE USEFULNESS OF THE MULTICRITERIA APPROACH FOR POLICY-MAKING

Multicriteria analysis has already shown its usefulness in the early stages of the process of policy-making. In the first place, the approach of multicriteria evaluation is a powerful tool in structuring the work of the planning group. Using this approach, the problem is divided into a number of clearly defined smaller problems (cf. Figure 1). This enables the planning group to more easily delegate some parts of its work to other research groups.

Furthermore, the nature of the method makes it possible to handle many properties (i.e. criteria) of the planning area at the same time in a manner which is easily grasped. In the case of the regional plan for Twente, up to sixteen criteria were used for one single function (namely, the housing function; cf. Figure 2). It can be concluded, with hindsight, that only four or five criteria for each function would have been enough to determine the final results, 90 per cent of the time.

Because of its clarity of arangement, the multicriteria methods are very convenient in generating alternative plans for evaluation. The most important places to make alternative assumptions are in the sets of weights. Moreover, all other exogeneous parts of the method implemented for Twente (cf. Figure 1) may be used as entries for alternative urbanization models. It is worth noting that in this manner, the multicriteria approach is used both to evaluate alternative locations, and to generate alternative allocations. In this case alternative locations to be presented for public discussion and subsequent decision-making, were not selected by means of a multicriteria evaluation. A sensitivity analysis was used instead.

In addition to its structuring effect, the multicriteria approach also has an objectifying effect, thus increasing our insight into the problem. Although some of our ideas may be confirmed, this effect can be best described as a dissolution of fixed ideas regarding the urbanization of the northern part of the Central Urban Zone of Twente.

Looking at the map of this urban zone (cf. Figure 4), the idea may arise (as it did) that the open area between Almelo and Vriezenveen is the most suitable site for new housing development, because of its proximity to the centre of Almelo, to good infrastructure, and to working areas. However, none of the explored alternative models suggested allocating the housing function to that area. Further examination of the data made it clear that these obvious thoughts neglected, or at least under-restimated, both the agricultural properties of the area in question and the absence of such qualities north of Wierden, as well as the housing properties of the area north of Wierden. Nevertheless, during the consultations, an alternative location for housing south of Vriezenveen has been added. This happened at the instigation of the representative of the national department of physical planning, for the reasons described above. Subsequent decision-making makes it seem that, in spite of everything, this fixed idea has been able to overrule objective arguments.

In the same way, the method worked as an eye-opener with respect to the southern development of Wierden. For similar reasons, this development, for many years favoured by the municipality of Wierden, can be considered less desirable than the northern development. Gradually, all other parties involved agreed on this point.

The third example of the teaching effect concerns the frequent allocation, in the

model, of the housing function south-east of Almelo and north of the railway (cf. Figure 7). On the neighbouring industrial park, Ultracentrifuge Nederland is located, a plant for the enrichment of uranium. As far as is known no danger zone or the like is in force for this plant. Nevertheless, no one would seriously propose this location for housing. It seems that phychological factors, like fear of nuclear accidents, may be more important for decisions about urban development than housing criteria are.

Finally, the multicriteria approach enables one to deal with very different functions in a similar way. In the context of conflicting interests between urban and rural functions, this property of the method proved to be decisive in convincing the representatives of rural interests, and especially those of agriculture interests, to give their opinion upon the locations that might be considered for sacrifice to urban development.

From a phychological point of view, one can say that an auspicious climate has been created for consultation and decision-making with regard to the urbanization in Twente.

REFERENCES

(1) Provincie Overijssel/Gewest Twente, Research paper 017, Province of Overijssel, Zwolle (1982)

(2) Provincie Overijssel/Gewest Twente, Voorontwerp herziening streekplan Twente, Province of Overijssel, Zwolle (1982)

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

STUDYING GRAVEL EXTRACTION THROUGH MULTICRITERIA ANALYSIS

S.J. Bennema (1)
H.'t Hoen (2)
A. van Setten (3)
H. Voogd (1)
(1) Dept. of Civil Planning, University of Technology Delft
(2) Province of Gelderland
(3) Research Centre for Physical Planning TNO

1. INTRODUCTION

The aim of this paper is to discuss some practical experiences the authors have obtained with the application of multicriteria techniques to a location analysis of a politically sensitive subject, i.e. the future gravel production in the Dutch Province of Limburg. It concerns a joint research project of the Research Centre for Physical Planning TNO and the Civil Planning Group of the Delft University of Technology, which has been commissioned by the Provincial Government of Limburg.

The study bears some interesting characteristics. For instance, it includes practical applications of recently developed multicriteria techniques on a political issue fraught with conflict. The various participants involved (viz., provincial politicians and civil servants) did not have actual experience with systematic project evaluation techniques which implied that during the study much attention had to be given to matters of communication and presentation. Besides, the researchers had to cope with inadequecies and gaps in the data and (concealed) differences of opinion amongst the participants about the ultimate objectives of the study.

In this paper the attention will be mainly focussed on methodological experiences and findings (see also Bennema c.s., 1985). It should be emphasized that the views and opinions being presented are those of the authors alone and that they do not necessarily represent those of the Provincial Authorities of Limburg or any other organization involved.

The structure of the paper is as follows. First, the planning problem and the objective of the study will be outlined in section two. In addition, the structure of the study is discussed in section three. Section four is devoted to some practical issues and problems, which have been encountered during the study. Finally, in section five some summarizing remarks are given.

2. THE PROBLEM

The main issue dealt with in the study concerned the extraction of gravel in the province of Limburg. Gravel is an indispensable building material. It is used in extensive quantities for purposes of civil engineering and housing. The demand for gravel in The Netherlands can amount each year to up to 20 million tons. Almost ninety-five percent of the Dutch production of gravel - almost 14 million tons - is

won by private firms in one region: Middle Limburg (see Figure 1). To satisfy the Dutch demand for gravel it is expected that gravel extraction in this area has to be continued through the next decades (see also Ike and Luypers, 1982).

The impact of gravel extraction on the regional structure is considerable. Most of the former pits are lakes now and they fulfil an important recreation function. Before 1969 the production of gravel consisted mainly of small scale gravel pits in the winter bed of the river Meuse. In order to avoid a further uncontrolled deterioration of the landscape, the Provincial Government decided in 1969 to concentrate future gravel production in a limited number of areas. They entered into a contract with the joint gravel producers. According to this contract until 1990 the firms have to restrict their applications for new production sites to areas which are assigned by the provincial government.

Figure 1. **Middle Limburg**

The gravel producers also have to pay a certain amount of money for each ton of gravel into a special fund. By means of this fund the provincial government finances the reconstruction of gravel-pits.

In the years after 1969 a new big pit has been created near Panheel-Beegden, an area outside the winter bed of the Meuse. Local citizens, municipalities and environmental organizations protested vigorously against this project. Major objections are, among others, the immanent isolation of small villages which are more or less permanently surrounded by open water, the emphasis in the pit-restoration plans on water recreation, and the disappearance of the characteristic landscape. Last but not least, critics are focussing on the loss of environmental and agricultural qualities in the surroundings of the gravel-pit due to hydrological changes caused by the gravel production.

The foregoing illustrates that various conflicting interests are involved in the planning of new gravel-pits, both local and national, finding expression in several formal and informal pressure groups. The major interests can be classified through the activities which may be affected by gravel production, viz. gravel-production itself, agriculture, forestry, environmental preservation and control, recreation, employment and housing.

The continuation of the present gravel extraction activities is assured until 1990. Since the procedure for opening an entire new quarry will take at least four or five years - due to many legal arrangements, the treatment of objections, etc. - the provincial authorities have already made a start with the preparations for the planning after 1990. This is why they decided in 1980 to initiate the research project, called EVOLIM, which is discussed in this paper.

The main objective of the EVOLIM project is to "gain insight into possible alternative policies and their consequences concerning the problem of gravel production in Limburg, such that the provincial government can make an accountable choice for the best policy to pursue in the future". For a good understanding of the study it is important to note that the prime purpose was **not** to designate the best possible location(s) for future gravel pits, but instead to provide the provincial government with information about the relative qualitaties of the various possible sites for future gravel production. In other words, it should not necessarily result in a recommendation where to produce gravel after 1990. We will come back to this issue in section four.

3. THE STRUCTURE OF THE STUDY

The study can be characterized as a step-wise "top-down" approach, whereby the study area is reduced in a systematic way such that the attention is finally focussed on a limited number of potential sites. The approach was such that in each step choices were made for which explicit - and hence accountable - reasons have been given. The various steps are roughly visualized in Figure 2.

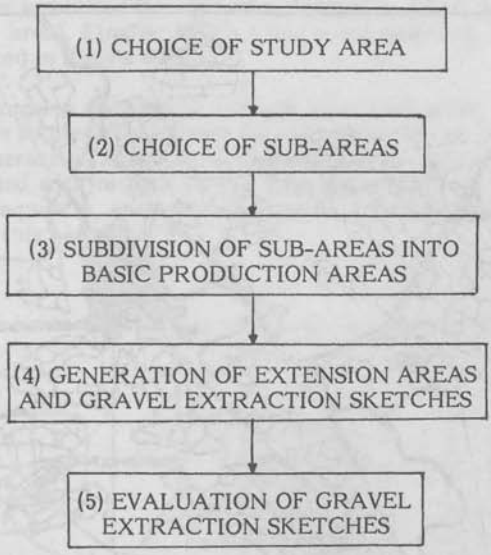


Figure 2. **Rough Outline of the Approach Followed**

Step one appeared in retrospect to be one of the most crucial and regrettable decisions in the study. It was decided by the responsible political executives that the study should focus on the Middle Limburg area, in particular on the area which was included in the official regional plan for North and Middle Limburg. Unfortunately, because of this decision some important areas suitable for gravel production in the southern part of Limburg were excluded. Although at first, this was not considered to be very important, the opinion soon changed due to the fact

that the gravel producers began buying land in this area.

In step two a socalled "sieve analysis" was applied, which simply means that areas were excluded from further consideration which did not meet a number of explicitly defined standards. Those areas were dropped which (a) are designated in the regional plan as "very valuable natural areas", (b) have urban or industrial functions, (c) include express ways or canals, and (d) are already designated to be used for gravel production in the near future (i.e. before 1990). The remaining subareas vary considerably in size (see Figure 3). Consequently, in step three a subdivision has been made into a large number of relatively homogeneous zones, which are called **basic production areas** (or **BPA's**). The size of each basic production area (app. 75 ha.) is such that in theory the area could be solely exploited. However, in practice always combinations of these areas will be made, in order to reduce the costs involved in the digging and transportation of gravel.

The BPA's can be combined in various ways, depending on their suitability for gravel production, their regional situation and the political view (i.e. interest(s)) that has to be emphasized. In step four various combinations of BPA's were made, which became the subject of a mixed (qualitative-quantitative) data multicriteria evaluation in step five (see for technical details about this evaluation method: Voogd, 1983)

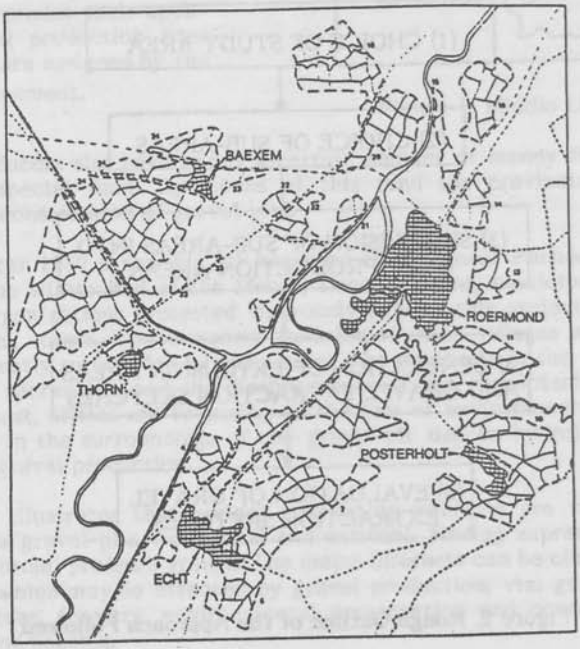


Figure 3. **The Basic Production Areas**

In Figure 3 a map is given of the basic production areas. These BPA's have been combined into **extension areas** (i.e. combinations of BPA's) and **gravel production sketches** (i.e. combinations of extension areas that meet an a priori defined standard with respect to the amount of gravel that must be extracted). The approach followed can be best illustrated through a small example.

In Figure 4 a hypothetical sub-area with six BPA's is given. Three of these BPA's are situated close to a waterway, so they are very suitable from a transportation point of view. However, first an overall suitability score is calculated for each BPA by means of a simple weighted summation approach. Because of the large number of BPA's, a more sophisticated method utilizing pairwise comparisons would have been less appropriate (i.e. too costly and unmanageable due to the large number of comparisons). The suitability scores are visualized in Figure 4-B. It is assumed that a "lower" score implies a "higher" suitability for gravel extraction.

The next step in the procedure is to focus on the BPA with the lowest score, which is c with score 2. By postulating that the gravel production starts here, the following step is to search for adjacent BPA's with an average score as close as possible to 2. These adjacent BPA's are b, d, e, g and h with an average score of 13/5 = 2.6. The addition of any other BPA would have resulted in a higher average value. The five BPA's together are called the **extension area** of c. Next, f can be added (i.e. the next extension area). Finally, BPA's i and a are selected. The various extension areas are ranked in Figure 4-C.

Evidently, each extension area has an average suitability score, which implies that the same procedure as illustrated above for the sub-areas can be repeated to generate the gravel extraction sketches for the entire study area. However, this time a constraint is added with respect to the minimum amount of gravel which must be produced. Consequently, each sketch includes as many extension areas as are necessary to meet this condition.

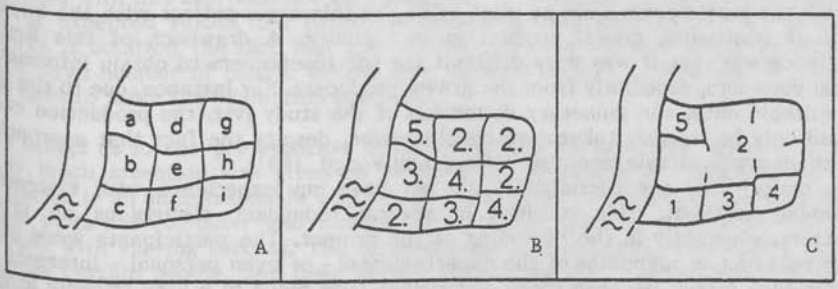


Figure 4. **The Selection of the Extension Areas**

The suitability scores mentioned before are based on a multicriteria analysis of the BPA's. For this analysis seven different weight sets have been defined, all stressing a different **point of view.** Besides, the depth of a pit is also taken into account. Evidently, the more gravel that is extracted, the more costly a pit restoration will be. Three different **extraction levels** have been distinguished, i.e. 100 ton per ha., 120 ton per ha. and 140 ton per ha.. Finally, explicit consideration is given to the amount of gravel that must be produced. Consequently, three socalled **gravel output categories** have been defined, i.e. a 175 million ton level, a 125

million ton level and a 75 million ton level. Consequently, in total 7 x 3 x 3 = 63 alternative gravel extraction sketches were generated.

All these sketches were evaluated by means of the three "mixed data" multi-criteria techniques, which are discussed in detail in Voogd(1983). Thirty-seven different criteria were formulated for this purpose. In addition, six qualitative (ordinal) weight sets were used, each stressing a particular priority view, i.e. a housing view, an agriculture view, an employment view, a recreation view, an environmental view and a forestry view. For each gravel output category and priority view an evaluation was performed of the alternative sketches. This meant that a large number of different rankings of the sketches were produced. By using a special condensation and presentation method, which is treated in the next section, a final recommendation could be given concerning the relative suitability of the areas in Middle Limburg for future gravel production. To support this task, a socalled **"gravel atlas"** has been created, which supplies the civil servants and other policy-makers with information upon which to base their further plan-making.

4. APPLICATION ISSUES AND EXPERIENCES

The project was guided by two different committees: a **working group**, which consisted mainly of practising planners and civil servants of various governmental provincial and national bodies, and a **guiding committee**, consisting mainly of chief-executives. The working group was chaired by the head of the research group of the Provincial Planning Department, whereas the guiding group was chaired by the political executive who had the first responsibility for regional planning and development. So, with exception of three political executives in the guiding committee and representatives of the research organizations, all other members of both groups were provincial officials or national officials with a provincial task. Non-governmental interest groups were not welcome to participate in either of the two committees, not even the gravel producers. Only in the final stage of the study the local governments were invited to participate. Consequently, the study was carried out in a relatively "quiet environment" in which the participants were at least willing to tolerate a project with the implicit aim of continuing gravel production in Limburg. A drawback of this limited audience was that it was very difficult for the researchers to obtain information from outsiders, especially from the gravel producers. For instance, due to the lack of suitable data, the monetary dimension of the study (viz. the production costs) could only be roughly taken into consideration, despite the fact that appropriate methods are available (see also 't Hoen and Voogd, 1981).
The majority of the participants did not have any experience with systematic planning methods. This resulted in several redundant discussions on trivial matters, especially in the beginning of the project. The participants were often only reacting as advocates of the departemental - or even personal - interests for which they stood. Because these discussions took place in a very friendly atmosphere, i.e. supporting one's owns interests without critizing other interests, the researchers often reached the conclusion that in the study area "everything is unique, irreplaceable, or at least extremely important".

Consequently, the sieve analysis took a lot of time. However, it was possible to reach an agreement by using the appraisals given in the (draft) official regional plan. The subdivision of the various sub-areas into basic production areas (BPA's) did not lead to much discussion. The criteria, on the other hand, were much more difficult to define.

In the first instance, the discussions in the working group were used by the researchers to gain insight into the issues they had to deal with. However, a point of departure for the study has been that no additional data-gathering could take place; i.e. the study should be performed with the available data. Unfortunately, there were hardly any suitable data available or only in a rough form on maps with a very small scale. It took a few months to arrive at a set of criteria which were both relevant and measureable in some way or another.

Despite the fact that the mixed data evaluation methods are able to deal with purely qualitative (ordinal) criterion scores, it was attempted to find quantitative indicators for all criteria. The reason behind this is that ordinal rankings must be justified. The best way to explain a ranking is by relating it to a substitute variable or set of variables which can be measured. For instance, a qualitative criterion such as "noise nuisance" can be approximated by counting the number of houses in relation to their distance from a gravel pit. The resulting quantitative numbers can be used to construct a ranking.

However, neither the criterion scores nor the weighting schemes (priority views), which were both proposed by the researchers, did evoke much debate. This may be explained by the fact that during the study no attempts were made to favour certain areas to accommodate future gravel production. Obviously, if such a common sense preference would have been available, no large study with external consultants would have been initiated. Besides, during the study the researchers continuously drew attention to the variety of possible priority views. As a consequence, a large number of alternative weighting schemes were used, not as empirically derived priority statements but as hypothetical assumptions for which others could be substituted. Obviously, this convinced the participants that they were not tied for ever to these priorities. Besides, in the evaluation phase only qualitative priority statements were used.
There are at least two satisfactory ways to deal in a multicriteria evaluation with ordinal weights, viz. through a random approach or by directly calculating the metric expected values of the ordinal weights (e.g. see Nijkamp c.s., 1984). In this study the random weight approach has been used (see also Voogd, 1983). The experiences with this approach are very pleasant. In contrast to the direct calculation of expected values, the random approach is very easy to explain to non-experts. In addition, the fact that this method does not use one single weight set but a large number of metric weight sets, appears to give confidence to persons who have to judge the evaluation outcomes.

This brings us to the presentation of the intermediate and final results. In the study much attention was given to methods of presenting multidimensional information in an understandable way. In the final report both three- and twodimensional plots and computer and manual cartographic maps were utilized for this purpose with varying success. The three-dimensional pictures, like Figure 5, did impress certain people, but they were - except for this characteristic - quite useless in a practical sense. In Figure 6 an example is given of more useful pictures which have been used to present the information from the evaluation of the gravel extraction sketches.

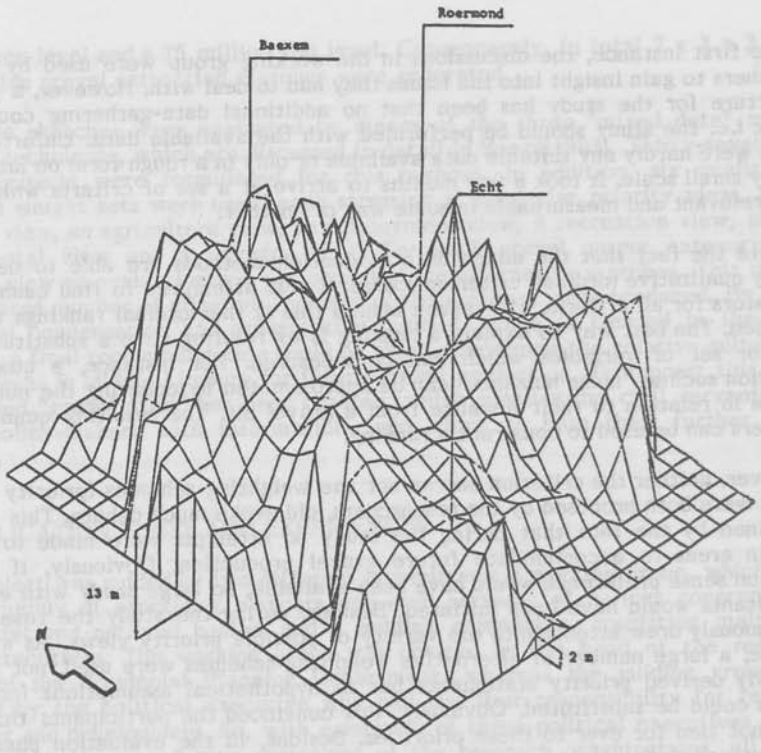


Figure 5. **A Three-dimensional Picture of Gravel Layers**

The alternative sketches are denoted horizontally, whereas the final ranking in the evaluation is represented vertically. The outcome of a certain run of an evaluation method is reflected by the position of the symbol for that particular run in diagram. If the various runs show identical outcomes for a given alternative, then the symbols are printed on the same position. Otherwise, there will be a vertical gap between the symbols, which reflects the uncertainty of the evaluation. Through these pictures a general conclusion can be drawn about the condensed final ranking. This can be illustrated by means of a small example: suppose four different runs have been made with the following results:

run 1 = A - C - D - B - E
run 2 = A - D - C - B - E
run 3 = A - C - D - E - B
run 4 = A - D - C - E - B

The symbol (-) means here "is preferred to ". These outcomes can be transformed into the following overall final ranking: A - (C,D) - (B,E), where (C,D) means that no choice can be made between these two alternatives.

Related to the presentation issue is the question about the use that must be made of the study. The "official" objective of the study has already been outlined in section two of this paper. It was also mentioned that the project did not have to result in a straightforward recommendation with respect to the "best" possible

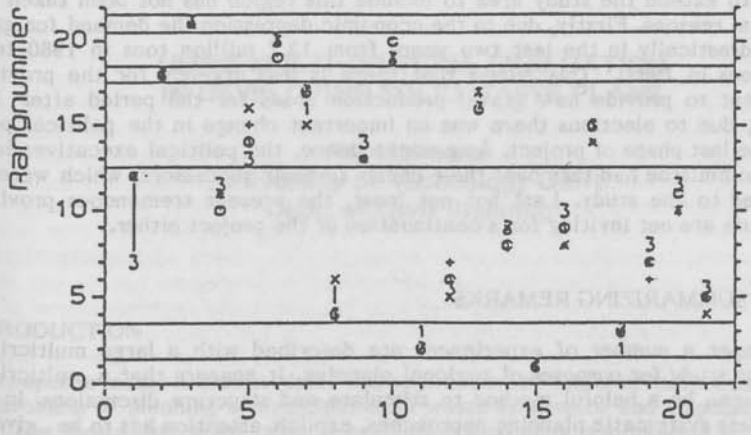


Figure 6. **An Example of the Pictural Presentation of the Results of the Sensitivity Analysis**

location. As a matter of fact, from the beginning, there was even an implicit fear amongst some of the participants that this might happen, thus leaving hardly any freedom for further negotiations and so forth. On the other hand, there were also some participants who implicitly seemed to prefer a straightforward result. This issue became subject of thorough discussions during step four (see Figure 2) of the study due to an increased insight into the relative empirical qualities of the BPA's.

Already the first efforts to combine the BPA's into larger units showed that there is hardly any area left in Middle Limburg, which is - at least moderate - suitable for gravel production for **all** priority views that were analyzed. On the contrary: the first pictures proved the highly conflicting nature of further gravel production in this region: areas which did score very well for some weighting schemes were completely unattractive from other points of view and vice versa. The preliminary outcomes also indicated that only a very small area of a few BPA's might be acceptable for all priority views under consideration.

These preliminary intermediate results evoked the decision to change the research strategy a little bit. Instead of throwing light only on an area's absolute suitability for gravel production, the question became now to highlight the areas which should become available if the region must follow the demand for gravel. This resulted in the introduction of gravel output categories and the approach outlined in the preceding section. Evidently, in order to meet the constraint with respect to the amount of gravel, all sketches now included areas which were hardly suitable for gravel production from certain priority viewpoints. It is obvious that these outcomes were not praised by everyone. Especially the environmental experts showed some dissatisfaction: an area with relatively high environmental values but also with a - from a geological perspective - rich gravel occurrence often entered a sketch because of this last feature.

Because of these results the initial decision to confine the study area to Middle Limburg is more and more questioned. It is evident that if South Limburg is also taken into consideration, many unattractive BPA's in the present sketches might

be exchanged with better areas in the soutern part of Limburg. However, the decision to extend the study area to include this region has not been taken (yet) for several reasons. Firstly, due to the economic depression the demand for gravel dropped drastically in the last two years from 13.7 million tons in 1980 to 9.1 million tons in 1982. This means that there is less urgency for the provincial government to provide new gravel production areas for the period after 1990. Secondly, due to elections there was an important change in the political power during the last phase of project. As a consequence, the political executives in the guiding committee had turn over their chairs to their successors, which were less committed to the study. Last but not least, the present tremendous provincial budget cuts are not inviting for a continuation of the project either.

5. SOME SUMMARIZING REMARKS

In this paper a number of experiences are described with a large multicriteria evaluation study for purposes of regional planning. It appears that a multicriteria approach can be a helpful method to stimulate and structure discussions. In contrast to less systematic planning approaches, explicit attention has to be given to all dimensions of the problem due to the use of qualitative and quantitative criteria and alternative weighting schemes. As a consequence, the project has had an important learning effect.

Another conclusion is that technical planning methods and approaches are acceptable in planning practice, provided that the basic principles **and** the methodological weaknesses are outlined clearly and - if possible - taken into account in one way or another. Evidently, a good verbal and written presentation is extremely important (but also sometimes very difficult).

The general conclusion can be drawn that from a methodological point of view, a number of interesting findings have been obtained, which certainly help to improve this kind of planning research. However, this study also illustrates that the success or failure of an approach is strongly related to the persons who are directly involved in the project and the organizational and social circumstances in which a project functions. Although at this stage no further analysis of this issue can be given, the topic itself certainly merits further consideration in planning research.

REFERENCES

Bennema, S.J., H. 't Hoen, A. van Setten and H. Voogd, Multicriteria Evaluation for Regional Planning: Some Practical Experiences, **Papers of the Regional** Science Association (P.W.J. Batey, ed.) (1985) (forthcoming)

P. Ike and H. Luypers, **Grind in de toekomst**, Publikatiereeks Civiele Planologie nr. 6, Delft University of Technology (1982).

H. Voogd, **Multicriteria Evaluation for Urban and Regional Planning**, Pion, London (1983).

H. 't Hoen and H. Voogd, A Note on the Financial Assessment of Plans, **Planologisch Memorandum** 81-6, Dept. of Civil Engineering, Delft University of Technology (1981).

P. Nijkamp, P. Rietveld and H. Voogd, A Survey of Qualitative Multiple Criteria Choice Models, **Measuring the Unmeasurable** (P. Nijkamp, H. Leitner and N. Wrigley, eds), Martinus Nijhoff, The Hague (1984).

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

USING MULTI-CRITERIA EVALUATION IN DEVELOPING SOLID WASTE PLANS

M. Maimone
University of Technology Delft
Dept. of Civil Planning

INTRODUCTION

This paper presents a case study of the application of multi-criteria evaluation to the problem of planning a municipal solid waste treatment and disposal program. It will be shown that multi-criteria evaluation (hereafter abbreviated as MCE) can play an important role in all the phases of planning, and not just in the evaluation process itself. Despite the fact that some interesting results were obtained from the evaluation, the paper focuses on the usefulness of MCE in organizing and solving a complex planning problem, and not on the actual content of the evaluation results.

The paper is organized as follows. First, the problem is sketched and the major difficulties are described. Thereafter, a planning framework within which a complete evaluation can be carried out is developed. This is followed by a section dealing with the actual application of the MCE. In the final part, a number of general conclusions are drawn about the use of MCE in developing a solid waste management program.

MUNICIPAL SOLID WASTE MANAGEMENT: A PROBLEM IN PLANNING

During the past ten years, the treatment and disposal of municipal solid waste has received the increasing attention of local and provincial authorities in the Netherlands. To a degree, this is a result of public pressure to address environmental problems resulting from numerous 'incidents' involving dumps, sanitary landfills, and composting plants. It is also a logical result of the environmental revolution of the 60's and 70's. Unfortunately, the task of choosing the most efficient and desirable treatment system for solid waste has not been as easy as was anticipated. Planners are faced with three factors, all of which add to the complexity of the problem. The first factor is an increase in the number of feasible, technical alternatives. To the three traditional methods of solid waste treatment, sanitary landfills, incineration, and composting, have been added resource recovery plants, resource derived fuel, and source separation methods. Various combinations of the above methods are also possible, adding to the number of viable alternatives to be analyzed.

A second source of complexity is the growing scale of the treatment systems being planned. The burden of decision-making, which formerly rested with local authorities, has shifted with the passage of the Solid Waste Law (Afvalstoffenwet, 1977) to the provincial governments. They have now been assigned the task of producing regional solid waste plans for the entire province, instead of allowing

each township to solve its own problems on a local level.

A third factor is an increasing awareness of externalities, both social and environmental. It is no longer considered acceptable to simply weigh financial costs and benefits, investigate the physical constraints, and choose the optimal alternative. Environmental costs, social acceptability, national objectives for hygiene, political and legal aspects, all demand 'equal time' with financial and technical considerations. For this reason, newer methods of evaluating alternatives need to be applied, methods which are flexible, easy to work with, and able to handle the complete array of evaluative criteria.

Upon examining the various solid waste plans that have been published to date, it is clear that a well defined planning and evaluation procedure has as yet to be applied (see Lauiszoon, 1982). Most decisions regarding the regionalization of the treatment of solid waste have simply resulted in the expansion or continuation of existing facilities. Despite a great deal of attention to the environmental aspects of solid waste treatment in the stated objectives of the various plans, these aspects either disappear when the actual decision is made, or are only made use of in a very rudimentary way. Multi-criteria evaluation techniques offer a way in which to include these aspects in an explicit and verifiable evaluation.

THE PLANNING PROCEDURE

Developing a collection, treatment, and disposal system for municipal solid waste demands that a number of separate but interrelated design decisions be made. The major decisions can be summarized as follows:

- methods to control or reduce the production of solid wastes by means of packaging and bottling laws, changes in production processes etc.
- methods of collecting the solid wastes. This is primarily a choice between combined collection in which all the solid waste is collected mixed, and source separation, where paper and glass are collected separately for recycling.
- a choice of transport systems (boat, rail, truck), as well as a collection plan coupling waste production points to treatment facility locations.
- the choice of an optimal treatment technique for the given region.
- the selection of a location for the chosen treatment system.

Ideally, all possible combinations of the above elements would be investigated and evaluated using a complete set of economic, efficiency, environmental, and degree of recycling criteria. This is, however, not possible due to the enormous number of combinations of design elements. It is also a rather mechanistic and inefficient approach, since partial evaluation of each element or decision can provide sufficient insight to narrow the choice of possibilities down to a more modest number.

It is not the place here to delve deeply into the actual process of developing the planning procedure. It is of importance to note that this developed procedure is predicated upon a number of premisses. First, the procedure is very much decision oriented. This means that it is assumed that the problem is clearly defined, and that the alternatives have already been developed. These are valid assumptions in the case of solid waste treatment, since there is a consensus on the urgency and definition of the problem, and the technical constraints are such as to limit the viable alternatives to a manageable number. By concentrating only on the most important decisions, the planning procedure can be designed as a linear decision sequence.

Figure 1 is a flow chart of this sequence. It consists of only the two **major** decisions that need to be made when developing a solid waste management program. These two decisions are answers to the questions "Which treatment system is best for the planning region?", and "What is the best location for this system?". The other decisions, relating to the transport system design and to the development of schemes to limit the production of solid waste, are left out of the linear decision model. The first, because it is fully dependent on the location of waste production centers and treatment sites, and can be dealt with only after the second stage of the procedure has been completed, and the second because it falls prior to or parallel with the planning method developed here, and is relatively independent of that procedure.

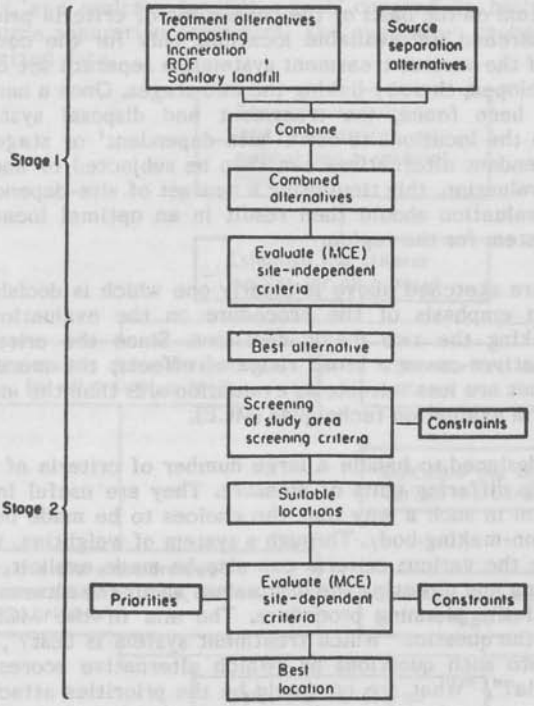


Figure 1. **Two-stage evaluation procedure for use in solid waste planning**

The linear decision model falls into two stages. During stage 1, the available treatment and disposal techniques are examined and evaluated as to their reliability. Only proven, viable techniques are considered further. The same sub-evaluation is applied to the available source separation techniques. Having thus screened out techniques which have not yet proven reliable in practice, the fol-

lowing step is to couple the source separation techniques to the treatment and disposal systems, producing the 'combined alternatives' which can then be evaluated using a mixed data multi-criteria evaluation technique.
The main idea behind the stage 1 evaluation is to provide a technical/environmental/economic evaluation, carried out on treatment and disposal systems completely independently of the future location of the system. It serves to establish a ranking of the alternatives, based on a given set of criteria priorities, so that only the best alternative need be carried over into the second stage evaluation. By isolating the evaluation of the various treatment and disposal methods from the effects of local circumstances, the results can be easily applied to a wide range of situations. This means that the evaluation does not need redesigning each time the local conditions vary. A simple adjustment of priorities and constraints is sufficient.

The stage 2 evaluation consists of two separate steps. In the first step, the planning area is searched for possible locations for the treatment system which proved to be the most desirable during the stage 1 evaluation. Thus, if composting proved to be the optimal system on the basis of the chosen set of criteria priorities, then the area would be searched for available locations only for the composting alternative. For each of the various treatment systems, a separate set of locational criteria has been developed, thereby linking the two stages. Once a number of possible locations have been found, the treatment and disposal system deemed optimal is coupled to the locations to form 'site-dependent' or stage 2 alternatives. These site-dependent alternatives can then be subjected to another mixed data multi-criteria evaluation, this time using a new set of site-dependent evaluation criteria. This evaluation should then result in an optimal location for the optimal treatment system for the region.

The planning procedure sketched above is clearly one which is decision oriented. This places the main emphasis of the procedure on the evaluation technique chosen to aid in making the two major decisions. Since the criteria used in evaluating the alternatives cover a broad range of effects, the more traditional cost-benefit approaches are less suitable as evaluation aids than the more recently developed multicriteria evaluation techniques (MCE).

MCE techniques are designed to handle a large number of criteria of varying importance, expressed in differing units of measure. They are useful in classifying the needed information in such a way that the choices to be made become much clearer for the decision-making body. Through a system of weighting, the political priorities attached to the various criteria can also be made explicit, which is of great use in stimulating and directing the discussions about the alternatives during the decision phase of the planning procedure. The aim of the MCE is not to provide an answer to the question "Which treatment system is best?", but to help gain added insight into such questions as "Which alternative scores higher for certain sets of criteria?", "What are or should be the priorities attached to each criterion?", and "Are certain alternatives **clearly** better than others?". Of paramount importance is the fact that the entire procedure is explicit and thus open to verification.

The MCE technique applied in this study is described in detail in Voogd (1983). For the purpose of this paper, it is only important to understand that the technique uses qualitatively and quantitatively weighted criteria to convert the evaluation matrix into a set of appraisal scores, which give an indication of the final ranking or worth of the alternatives under consideration. The evaluation matrix, a cross table of alternatives and criteria, can have both ordinal and cardinal values within the single matrix. The final ranking is, of course, predicated on the assumptions

which lie behind the chosen criteria, the priorities used for these criteria, as well as the assumptions underlying the evaluation method itself. These assumptions can lead to uncertainties about the validity of the results, uncertainties which must be taken into account in analyzing the results.

The remainder of this paper concentrates only on the application of the MCE technique to the stage 1 evaluation, in which a set of eight possible treatment and disposal systems are compared to each other. In this way, the effectiveness of the technique in aiding the decision-making process can be demonstrated.

THE EVALUATION PROCESS

As mentioned above, the stage 1 evaluation was performed on the eight alternative treatment and disposal methods found to be viable during a preliminary screening procedure. These eight alternatives consist of the four major, proven, treatment techniques: incineration with heat recovery, composting, resource derived fuels, and sanitary landfills, each coupled to both a maximum and a minimum source separation procedure. The evaluation process itself was carried out using a mixed data

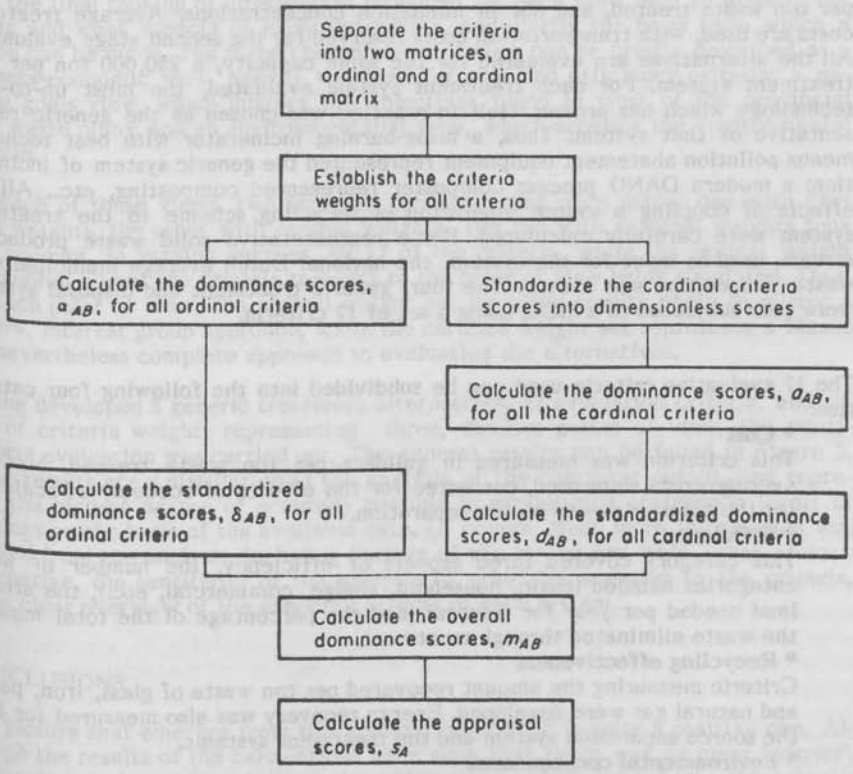


Figure 2. **Flow chart for the EVAMIX program**

multi-criteria evaluation program called 'EVAMIX', a program described in detail in Voogd (1983). A flow chart of this program can be found in Figure 2, and is, for

the purposes of this paper, sufficiently self explanatory. In this section, the emphasis is given to only two aspects of the evaluation procedure, the choice of criteria, and the problem of establishing a weight set to represent tahe relative importance attached to these criteria. These two aspects are, in general, most crucial, and to a large extent determine the effectiveness of the evaluation.

Developing a set of criteria with which to evaluate alternatives is always a source of uncertainty in any evaluation. This uncertainty, though difficult to control, must be minimized. This can be done by paying attention to the following questions, "Is the set of criteria chosen complete (that is, fully representative of the concerns of the interest groups involved) ?", "Do the criteria fully represent and describe the aspect for which they are intended as a measure?", and "Are the criteria evenly distributed over the various relevant aspects?". Such questions are always open to debate. The set of criteria chosen for this analysis were in part developed on the basis of the information available, and in part developed with the intenrion of provinding as well balanced mix of criteria. The criteria are specifically designed for a site-independent, technical evaluation, and cover environmental effects, cost, efficiency, and the degree of recycling achieved. Every effort is made to minimize the effects which specific, local circumstances might have on the values found. Thus, environmental criteria are expressed in emissions per ton waste treated, and not in immission concentrations. Average treatment costs are used, with transportation costs reserved for the second stage evaluation. All the alternatives are evaluated for the same capacity, a 250,000 ton per year treatment system. For each treatment system evaluated, the most up-to-date technology which has proven itself in practice was chosen as the generic representative of that system. Thus, a mass-burning incinerator with best technical means pollution abatement equipment represented the generic system of incineration; a modern DANO process composter represented composting, etc.. All the effects of coupling a source separation or recycling scheme to the treatment system were carefully calculated. For a representative solid waste production stream, used as input for the system, the national Dutch average municipal solid waste composition was chosen. The four, generic treatment and disposal systems were then subjected to a MCE, using a set of 17 criteria.

The 17 evaluative criteria used can be subdivided into the following four categories:

*** Cost**

This criterion was measured in guilders per ton waste treated. National average costs were used, corrected for the effects of economy of scale and for the changes due to source separation.

*** Efficiency**

This category covered three aspects of efficiency, the number of waste categories handled (toxic, household, sludge, commercial, etc.), the area of land needed per year for disposal, and the percentage of the total mass of the waste eliminated through treatment.

*** Recycling effectiveness**

Criteria measuring the amount recovered per ton waste of glass, iron, paper, and natural gas were developed. Energy recovery was also measured for both the source separation system and the treatment systems.

*** Environmental consequences**

This set of criteria was the most difficult to develop, primarily because of the theoretical problems involved in measuring effects in a site-independent fashion. The criteria developed were designed to evaluate the types of toxic substances emitted, the expected effects due to non-toxic emissions, the receiving medium for the toxic emissions (air, water, soil), the expected

average concentrations of toxic substances relative to the legally established allowable concentrations, the area affected by the emissions, the chance of human exposure to toxic substances, and a measure of the environmental impact of the source separtation system.

For most of the criteria, cardinal values were developed, however the environmental criteria were later 'downgraded' to qualitative criteria to reduce the uncertainty surrounding the accuracy of the values found. This uncertainty is due to the possible site-specific differences in effect once the precise location of the treatment system has been established. This particular uncertainty is much greater for the environmental criteria than it is for the others.

Developing a set of criteria priorities represents another kind of difficulty than does the problem of choosing the evaluation criteria. When making decisions in the public sector, where accountability is considered important, priorities which can affect the outcome of a decision should ideally be open and explicit. In this ideal situation, total consensus over priorities is reached and an exact representation of these priorities in cardinal weights is then established. This situation is rarely feasible. In this study, the problem was solved by creating three distinct 'points of view', in which artificially extreme weight sets were assigned to each view. In this way, the final ranking of alternatives produced for each view help to amplify the strong and weak aspects of each alternative, and clarify the basis upon which a choice can be made. The three chosen points of view can be briefly described as a business-economic view, heavily weighted for cost and efficiency criteria; a national goals view, which mirrors the objectives found in most of the provincial solid waste plans; and an environmentalist view, which stresses the environmental criteria.

For each of these views, two separate computer runs were made, one with cardinal weights, the other with ordinal weights. The run with cardinal weights took into account, in varying degrees, all 17 of the criteria, while the ordinal weight set ignored all but the criteria relevant to the point of view being simulated. Thus, for each point of view, the ordinal weight set produced results likely to mirror a narrow, interest group approach, while the cardinal weight set represents a biased but nevertheless complete approach to evaluating the alternatives.

Having developed 8 generic treatment alternatives, 17 evaluation criteria, and six sets of criteria weights representing three, distinct points of view, the multi-criteria evaluation was carried out. The general results can be found in Figure 3. These results are a distillation of the analysis of the computer output, and represent the highest degree of accuracy possible in establishing preference levels or rankings on the basis of the available data. Of course, much more information was derived from the analysis, including listings of the strong and weak areas of each alternative, the sensitivity of the analysis to priorities assigned to the criteria, and a clear overview of the expected effects of each system.

CONCLUSIONS

The picture that emerges from this application of MCE is quite a positive one. Although the results of the calculations were such that there was no clear 'winner', each point of view did produce results which could be sub-divided into three levels of preferability. More important than the final rankings, however, was the organizing influence of MCE on the entire planning procedure. In this final section, a number of conclusions regarding the usefulness of the two-stage planning procedure and the effectiveness of MCE within that procedure are summarized.

Preference level	Point of view		
	Business-economic	National goals	Environmental
Most desirable alternatives	CP+ CP-	CP+ CP- IN+	CP+ SL+ IN+
Middle group	IN+ IN- SL+ SL-	- -	IN- CP- SL- RDF+
Least desirable alternatives	RDF+ RDF-	IN- RDF+ RDF- SL+ SL-	RDF-

Abbreviations used: CP = composting, IN = incineration, SL = sanitary landfill, RDF = refuse derived fuel, + = maximum source separation, - = minimum source separation.

Figure 3. **The resulting preference groupings per point of view**

The two stage evalution procedure, in which the various alternative treatment sytems are first evaluated in a general way, followed by an evaluation procedure designed to rank the alternative locations for each system had the important advantage of clarity. Both of these decisions are well defined in subject and in evaluative criteria. The consequences of a decison in favor of composting at one location, for example, can easily be checked against other possible systems, planned at **another** location, without the complicating influence of site-specific locational effects interfering with the results. Thus, a 'cleaner', more technical decision can be made on the economic, environmental, and efficiency aspects, reserving the more politically charged, local considerations to the second stage, when a site for the treatment system must be selected.

Another advantage of the two-stage procedure is that the general characteristics and effects of a particular system, effects which will invariably arise, are clearly highlighted. This is not only an aid in keeping the discussion during stage 1 as rational as possible, but also aids in developing mitigating measures to reduce unwanted effects. Having separated the effects into a technical and a spatial category, technical improvements and land use planning can more easily be devised to directly reduce these effects.

The major weakness of the two-stage procedure is conceptual. The division into site-dependent and site-independent evaluations is in some ways artificial, as overlap is bound to occur between the two stages. In particular, an attempt at environmental assessment, free of local conditions, is a theoretically difficult problem, and much work remains in improving the reliability of the findings and in devising a suitable set of criteria. Only a first step in solving these problems has been made in this study. Another difficulty arises from the, of necessity, general nature of the values for each criterion found in the first stage evaluation. A larger degree of uncertainty in the results may be introduced, since local conditions may well influence the criteria values in ways not easily forseen. Care must be taken to accept as significant only large differences in effects between alternatives.

Turning to the use of a multi-criteria evaluation technique in the planning process, the 'reviews' are to a large degree positive. The technique has a very favorable influence on the entire process of developing a planning procedure. In the early stages, it helps to focus attention on the major moments of decision. In this way, the inter-relationships of the various decisions are made clear, and a logical, linear, decision sequence can be devised. Having devised this sequence, the use of a MCE technique helps to direct the research and information gathering phase, since information must be sought only to assign specific values to the already developed evaluation criteria. In short, one knows exactly what to look for. Another advantage is the freedom to use the information on effects in its natural dimension. No information need be lost because of mathematical problems of comparability. Also, both cardinal and ordinal values can be handled, thus allowing previously unusable 'soft' information to be utilized alongside the 'hard' data. A final positive aspect of the MCE technique is the help it offers in devising clearly defined alternatives, evaluative criteria, and priorities. Because of the structure of the technique, all three aspects must be explicitly dealt with. This helps to eliminate the production of inexplicable results and to further rationalize the decision-making process.

The major weakness of the procedure is only a weakness if one is searching for a definitive answer to the question "Which system is best?". As can be seen in Figure 3, the results depend greatly on the priorities assigned to the various criteria. In effect, one can manipulate the chosen weight set to arrive at a desired result. To avoid this problem, the informative aspect of the technique must be stressed. One way to do this is to provide results for a number of points of view. This implies that MCE techniques must not be used to make decisions, they should only help to inform the decision makers of the consequences of their decision.

To summarize, MCE can play an important role in planning a solid waste treatment and disposal system. It provides a conceptual framework around which the entire planning process can be developed. From the earliest plan forming phase through the collection and analysis of information to the actual decision making phase, MCE is an extremely useful tool. With more practical experience in its use, and the development of more accurate data, MCE could well become standard procedure for developing solid waste management programs in the future.

REFERENCES

Lauiszoon, P., Vergelijkend Onderzoek naar Provinciale Afvalstoffenplannen ex. Art. 4 Afvalstoffenwet, **Tijdschrift voor Milieu en Recht**, 2, Febr., (1982)

Voogd, H., **Multicriteria Evaluation for Urban and Regional Planning**, Pion, London (1983)

CONCLUSION

In: A. Faludi and H. Voogd (eds)
Evaluation of Complex Policy Problems
Delftsche Uitgevers Maatschappij (1985)

EVALUATION OF COMPLEX POLICY PROBLEMS SOME CONCLUDING REMARKS

H. Voogd
A. Faludi

INTRODUCTION

This last section summarizes the conclusions drawn from the papers and the very lively discussions that took place during the workshop held at the Delft University of Technology in The Netherlands in February 1984.

First, attention goes to the issues about which the participants were more or less in agreement. Then we discuss a number of topics which are likely to remain the subject of continuing debate. These include the problem of attaching weights to the achievement of certain objectives, and the ever-present dilemma between the pursuit of methodological refinement on the one hand and transparancy of evaluation methods on the other. Following on from that, we summarize implications for planning practice and indicate new topics for research.

EVALUATION AND DECISION-MAKING

Virtually all discussions during the workshop turned on planning and decision-making. All evaluation aims at responsible decisions, and decisions can never count as responsible without some form of evaluation.

Many practical problems in evaluation relate to the **context** in which decision-making takes place. Several authors (Hill, Lichfield, Buit, Sorber, etc.) emphasized **variety**, both as regards this context, as well as regards the **content** of evaluation, or its object. For instance, it matters a great deal whether evaluation is performed in an organisational setting with many actively competing interests (actors, departments, and so forth) or whether it occurs in a rather less turbulent environment. The case-studies (see for instance Bennema **et al.**) illustrate the extent to which the organisational setting influences success and failure in evaluation. As regards content, locational problems, for instance, may require examination of a large number of alternative sites. In more conventional plan evaluation, on the other hand, the number of alternatives is usually smaller - although a systematic approach to the generation of alternatives like AIDA can result in very large numbers of alternatives as well. Where the problem is to evaluate such a wide range of alternatives, there methods based on systematic weighting come in their own. The same methods make little sense where only a handful of alternatives are involved. This only goes to show that there is no one evaluation method which is capable of dealing consistently with the variety of contexts and contents. In this respect an interesting remark has been made concerning the robustness of evaluation. In view of the dynamics of planning and decision-making processes much more attention should be paid to what may be called **adaptive evaluation.** It can be described as an approach which is flexible enough to deal with changing circumstances during the decision-making process. An elaboration of this idea can be found in Hickling's paper.

Systematic evaluation always includes at least the following two elements: First there is the **structuring of information** by means of some kind of evaluation matrix. Second there is the **interpretation** of this information so as to arrive at some conclusion - mostly in the form of a recommendation as to the type of decision which should be taken. This necessarily includes some form of aggregation, involving the reduction of information.

As long as the number of alternatives under consideration and the number of judgement criteria remains manageable, the more important of the two is undoubtedly the evaluation matrix. Experience with evaluation in practice suggests that there is more emphasis on evaluation methods as a way of illuminating the dimensions of a decision and the implied conflicts, rather than simply as a device for identifying the best alternative. This throws light on the application of more complicated methods to facilitate the interpretation of evaluation matrices. Their value for practical decision-taking is therefore limited to complex and comprehensive problems. Usually, it is sufficient to construct a simple goals-achievement (or project-effect, or impact) matrix. Sometimes, a more elaborate planning balance sheet (see Lichfield) may be constructed.

Both the contributions of De Vos and Van Staalduine and the resulting discussions lead to the conclusion that the use of systematic evaluation methods helps **structuring planning work.** The methods not only stimulate, guide and broaden discussions (hence avoiding the danger of too much emphasis being given to one single issue), they also assist in understanding the problem. In addition, by identifying those aspects and/or criteria which are both crucial and difficult to assess, evaluation makes a more efficient use of manpower resources possible.

People often enter a decision-making process with a preconceived opinion about the most desirable outcome. Evaluation methods not only force decision takers to critically assess their pet-solutions in the light of its alternatives, it also stimulates the search for alternatives which satisfy the interests of as many participants as possible. As a consequence, evaluation methods **may engender innovation** in the sense of leading to solutions which would otherwise not have emerged. An example can be found in the paper by Vos.

The existing literature pays scant attention to the relation between the strategic choice approach and systematic evaluation methods. This depite the fact that they have much in common. As Dello shows, **strategic choice and evaluation methods** are largely **complementary,** and both approaches may be very well combined. However, this may increase the complexity of each, especially where all stages of the decision-making process are treated systematically. Also, as Van de Graaf illustrates, the aim of the strategic choice approach may be to reach agreement on certain issues rather than to arrive at an explicit assessment of alternatives. Where this is the case, the latter may even be counterproductive. After all, demonstrating the existence of conflicting objectives does not help with achieving consensus.

THE CONTINUING DEBATE

During the workshop, several issues emerged which seem to form the object of continuing debate and are likely to do so in the future. Some of them are - and will probably always be - very controversial, for instance the issue alluded to above of whether a planner, or policy analyst, must work towards **consensus,** or whether he should only feel responsible for the **quality of the information** avail-

able to decision takers. Obviously, the practical conduct of planners will always reflect both considerations. Also, the balance of emphasis may depend on the context and content of evaluation and on the phase of the decision-making process. Although the question cannot, therefore, be answered in any unambiguous way, it certainly addresses a dilemma which those involved experience in their practical conduct.

In a sense we can compare this dilemma with the gap that exists between the philosophy of the strategic choice approach on the one hand and formal, "classic" decision theory and **ex-ante** evaluation methods on the other. As is illustrated by the contributions of Hickling and De Graaf, evaluation in strategic choice is a group process in which the prime objective is to find a solution which seems acceptable to those involved. However, this need not be the "best" solution - nor indeed an acceptable one - to outsiders. Classic decision theory, on the other hand, seems to be less afflicted by this problem. It focuses on the quality of the solution, implying optimization, or at least satisficing, in the light of some explicit, pre-defined standards. But the acceptability of the outcome of such exercises depends on how meaningfull these standards are. Very often, though, they are limited to quantifiable effects and in danger of mis-representing the problem.

Current thinking concerning **ex-ante** evaluation represents a mixture between the philosophy of strategic choice and formal decision theory. Various contributions in this volume demonstrate that there is a search taking place for the proper balance between a methodology emulating sophisticated - but sterile - formal decision theory on the one hand and the exigencies of public decision making in practice on the other. A recurring theme in this respect is the attachment of explicit weights to criterion scores. Especially where the number of criteria is small, often a recommendation about the ranking of alternatives can be given without attaching numerical weights to these scores. But where there is conflict between various criteria, there some kind of statement of priorities is a precondition of being able to arrive at a definite conclusion. The opponents of weighting, on the other hand, argue that such explicitness in weighting reduces the acceptability of the outcome of evaluation exercises to the client, one of the main reasons being that the weighting of criterion scores by experts seems to court the danger of their arrogating to themselves what is essentially a prerogative of political decision takers. Advocates of explicit weighting usually counter that methods which evade this problem also use (implicit) weights, but without rendering them explicit, and thus accessible to public scrutiny. Although no unambiguous conclusion has been drawn during the workshop, a general opinion was that, suggesting a spurious degree of precision as it does, detailed numerical weighting is too artificial to be trustworthy. Qualitative weighting (using verbal statements or presenting alternative priority statements on an ordinal scale) should be preferred.

The discussion above undoubtedly bears a close relationship to the issue above of whether evaluation should aim at (group) consensus rather than improving (public) accountability. The issue of accountability is especially important where the results of an evaluation exercise must be presented to outsiders (**external evaluation**). Where, on the other hand, evaluation is performed with no such purpose in mind (**internal evaluation**), fewer restrictions apply. Once again, the ultimate choice of the preferred approach is always a function of the nature of the problem, the interested parties, the decision-making context, and so forth (see for instance the papers by Hill and Sorber), and in this respect, therefore, no clear-cut recommendation can be given.

An important consideration in evaluation concerns the question "who gains and who loses". As Miller shows in his paper, the dilemma of equity versus efficiency

is very difficult to deal with. In practice, the very definition of what equity involves is problematic. The reason is that such a definition may have many political implications. The assessment of distributional effects of planning proposals implies that explicit consideration must be given to social categories (for instance, the definition of groups of actors involved in, or affected by, policy). This may be very difficult, especially where the evaluation is carried out in relative isolation and without close consultation, therefore, with the various groups concerned. Where the outcomes do not reflect the preferences of one of the groups, this group is certain to voice criticisms. Close cooperation with the interests concerned is strongly recommended, therefore. Such citizen participation is a special requirement in external evaluation. In these cases, though, it is inadvisable to make explicit references to the actual groups or organizations concerned. More general and fictitious designations such as "motorists", "environmental interests", "low-income groups", and so forth should be used instead. Otherwise, the outcomes might not be acceptable to the groups concerned.

Often, evaluation is seen (and perhaps even used) as a way of **rationalizing** decisions already taken. During the workshop the question was raised whether such practices should be condoned, or whether they should be rejected as representing too limited a view of the role of evaluation. After all, the client should be able to form a balanced judgement concerning the decision problem and possible solutions. This is certainly an issue for continuing debate (see also Voogd, 1985).

CONCLUSIONS FOR PLANNING PRACTICE

The various practical examples documented in this volume, and the discussions during the workshop make it possible to formulate a number of recommendations for planning practice.

The first lesson for planning practice is that, however important and interesting (read: difficult to solve) they may be, one should resist the temptation of raising methodological issues in a practical context. The focus should be on the problem at hand and **not** on the method. Experience teaches that methodological debates - for instance stemming from the fact that there is no single 'best' evaluation method - are time-consuming and not always amenable to satisfacty resolutions.

Another important lesson concerns the determinate influence which evaluation methods may have on the planning process, and in particular on planning research. It helps to determine the types of investigation that must be undertaken. By first performing a crude qualitative evaluation, insight can be gained as to how important certain aspects (e.g. costs, criteria) are to the outcome of the exercise. Where it transpires that an issue is of critical importance to the final outcome, there additional discussions and/or research may be needed.

The next recommendation for planning practice is to choose the method which suits constraints of time and money. Of course, this may not necessarily be the "best" method. But in the choice of evaluation methods, as elsewhere, the best can be the enemy of the good. So it is sometimes better to perform a simple evaluation than to attempt a complex one and land up with no evaluation at all. The danger of using complex approaches is well-known: it may take much time to collect the necessary information and the methods may be demanding in terms of skilled manpower. In addition, there is always the danger that a too sophisticated approach will prove inaccessible to the public at large, thus generating mistrust. However, a good presentation of the evaluation results may overcome this problem (see also Daru).

The last recommendation concerns the object, or content of evaluation, or that which is being compared. Too many applications merely compare the end-states envisaged after proposed plans or policies are carried out. They remain on the level of what the literature commonly describes as 'blueprint' planning. Both the dynamics of the problem at hand as well as the difficulties of implementing proposed alternatives are being neglected. Several participants of the workshop stressed the necessity of paying more attention, therefore, to issues of implementation, both in relation to costs and consequences.

NEW DIRECTIONS FOR RESEARCH

Above, a number of issues in need of further investigation have been raised. In addition, some general conclusions can be drawn with respect to further research in the field of **ex-ante** evaluation of plans and policies.

An important issue for future research concerns the development of practical tools for impact analysis. As has been outlined by Lichfield, the purpose of **ex-ante** evaluation is not only to describe and synthesize impacts into some kind of recommendation, but also to discover and assess impacts in the first instance. However, presently there is still a lack of operational methods of impact analysis, in particular methods which can cope with qualitative or "soft" information in a consistent way. In addition, more thinking is needed concerning the relationship between impact analysis and plan or program formulation. For example, an interesting new avenue to explore is the possible relation between systems reasoning and procedural reasoning on the one hand and impact analysis on the other hand.

A key element with respect to the credibility and acceptability of evaluation is the way in which both methods and results are presented. The importance of good presentation not only holds for the equity issue, as has been discussed before, but is evident in all kinds of evaluation. Due to modern technology, many improvements are already on the horizon with respect to the use of computer graphics, and so forth. However, this is only one part of the story. A great deal of information resulting from evaluation cannot be represented graphically. Until now, hardly any research has been done with respect to possible 'interfaces' between evaluation and actual decision making. More research is needed, both empirical (case studies) and theoretical, to cover this important area.

In the previous paragraph mention has been made already of modern technology. It is to expected that in the near future more and more attention will be paid to computer-assisted information management. This will undoubtedly affect policy making processes, and hence also evaluation. As Nijkamp shows, several new organisational and methodological concepts emerge, such as management information systems and decision support systems, which all bear some relation to evaluation. Further research into this is certainly necessary, if only to make sure that information systems also include "soft" information in order to be sufficiently useful in practice.

Finally, attention should be paid to legal and procedural arrangements with respect to evaluation. In several countries there is already (some) experience with legally required types of evaluation, such as environmental impact statements, or sunset legislation. An international comparison of planning systems as regards the role of such evaluation seems very desirable indeed.

Of course, there are issues for fundamental research also. The influence, signalled

above, of the context within which evaluation takes place, and the substantive content of the policy problem to which it refers, on the nature of, and the methods used in, evaluation raises the issue of whether "evaluation" always denotes one and the same activity. If we want to bring some kind of unity to the field, then we would have to conceive of decision situations of utmost complexity, and formulate, at least on a conceptual level, a matching approach to evaluation. All practical decision problems, together with the appropriate practical approaches to evaluation, would then be simplifications of these general concepts. Something akin to this idea is underlying the taxonomy as presented by Hill. Clearly, it would require considerably more thought to work out in detail.

Another topic of a similar nature relates to the ethical theory underlying evaluation. It was observed at the workshop that all evaluation methods discussed had one thing in common: the expected outcomes were seen to be of crucial importance in determining the desirability of action. This represents one type of ethical theory: consequentialism (see Regan, 1981). Attention was drawn to another ethical theory - deontology - which radically brakes with this assumption. According to it, features of the act itself, rather than its outcomes, determine whether it is justified. Frankena (1973) relates this to the equity issue in particular - which only goes to show that an exploration of ethical theories might indeed be relevant to the concerns discussed at this workshop.

REFERENCES

Frankena, W.K., **Ethics** (second edition) Prentice-Hall, Englewood Cliffs, N.J. (1973).

Regan, D., **Utilitarianism and Co-operation**, Clarendon Press, Oxford (1981).

Voogd, H., Prescriptive Analysis in Planning, **Environment and Planning B**, vol. 12 (1985) (forthcoming)

LIST OF CONTRIBUTORS

Ir. S.J. Bennema
Urban and Regional Planning Group
Dept. of Civil Engineering
University of Technology
P.O. Box 5048
2600 GA Delft
The Netherlands

Professor dr. J. Buit
Department of Social Geography and Planning
Free University
P.O. Box 7161
1007 MC Amsterdam
The Netherlands

Ing. P. Daru
Faculty of Architecture, Building and Planning
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands

Dr. P. Dello
Institute for Social and Economic Geography
University of Louvain
Celestijnenlaan 300
3030 Heverlee
Belgium

Professor dr. A. Faludi
Institute for Planning and Demography
University of Amsterdam
Jodenbreestraat 23
1011 NH Amsterdam
The Netherlands

Ir. R. van de Graaf
Ministry of Economic Affairs
Directorate-General for Industry
P.O. Box 20101
2501 EC The Hague
The Netherlands

A. Hickling
Allen Hickling and Associates
No. 2 The Old Bakery
Church Road
Long Itchington
Rugby CV23 8PW
Great Britain

Professor dr. M. Hill
Faculty of Architecture and Town Planning
Technion-Israel Institute of Technology
Haifa 32000
Israel

Ir. H. 't Hoen
Province of Gelderland
Physical Planning Agency
P.O. Box 9090
6800 GX Arnhem
The Netherlands

Professor dr. N. Lichfield
Nathaniel Lichfield and Partners
Star House
104/108 Grafton Road
London NW5 4BD
Great Britain

Ir. M. Maimone
c/o 55 Spruce St.
Roslyn Harbor NY 11576
United States

Professor dr. D.H. Miller
Department of Urban Planning
University of Washington
410 Gould Hall
Seattle, WA 98105
United States

Professor dr. P. Nijkamp
Economic Faculty
Free University
P.O. Box 7161
1007 MC Amsterdam
The Netherlands

Ir. A. van Setten
Research Centre for Physical Planning TNO
P.O. Box 45
2600 AA Delft
The Netherlands

Drs. A. Sorber
Department of Policy Analysis
Ministry of Finance
Casuariestraat 32
2500 VB The Hague
The Netherlands

Drs. J.A. van Staalduine
National Physical Planning Agency
Willem Witsenplein 6
2596 BK The Hague
The Netherlands

Ir. J.W. de Vos
Economic-Technological Agency North Holland
P.O. Box 3007
2001 DA Haarlem
The Netherlands

Dr. H. Voogd
Urban and Regional Planning Group
Dept. of Civil Engineering
University of Technology
P.O. Box 5048
2600 GA Delft
The Netherlands