Open access • Journal Article • DOI:10.1177/0193841X8601000101

The Development and Trials of a Decision-Making Model An Evaluation

Michael A. Peters, James D. Marshall, Robert Keith Shaw

Institutions: University of Auckland

Published on: 01 Feb 1986 - Evaluation Review (SAGE Publications)

Topics: Summative assessment, R-CAST, Formative assessment and Professional development

William H. Schmidt
Nick L. Smith
George H. Wolkon
James Wright
James Zeisel

—Richard A. Berk
—Howard E. Freeman

We describe an evaluation undertaken on contract for the New Zealand State Services Commission of a major project (the Administrative Decision-Making Skills Project*) designed to produce a model of administrative decision making and an associated teaching/learning package for use by government officers. It describes the evaluation of a* philosophical *model of decision making and the associated teaching/learning package in the setting of the New Zealand Public Service, where a deliberate attempt has been initiated to improve the quality of decision making, especially in relation to moral factors.*

THE DEVELOPMENT AND TRIALS OF A DECISION-MAKING MODEL

An Evaluation

MICHAEL PETERS
JAMES MARSHALL
University of Auckland

ROBERT SHAW
New Zealand State Services Commission

I. BACKGROUND

This article describes an evaluation undertaken on contract for the New Zealand State Services Commission of a major project (the Administrative Decision-Making Skills Project), designed to produce a model of administrative decision making and an associated teaching/learning package for use by government officers. The State Services Commission is one of the three control departments (along with the Treasury and the Ministry of Works and Development) for government

AUTHORS' NOTE: *We wish to acknowledge the assistance of both the New Zealand State Services Commission and two anonymous reviewers of* Evaluation Review. *The views expressed in this article are our own and are not necessarily those of the State Services Commission.*

EVALUATION REVIEW, Vol. 10 No. 1, February 1986 5-27

administration in New Zealand. It has overall responsibility for the management and organization of government departments and the training and welfare of government officers (as laid down in the *State Services Act 1962,* and as amended subsequently). The project originated with concern over the implementation of the *Official Information Act 1982* and the number of departmental decisions taken by officers that were being questioned by the ombudsman. Recently in Western democracies there have been calls for greater open government, including the demand for greater public access to official information. The New Zealand *Official Information Act 1982* follows similar statutes passed in the United States (*Freedom of Information Act 1966,* amended 1974), Denmark, Norway, Canada (at the federal level) and Australia. It has been regarded as the most important constitutional innovation in New Zealand since the office of ombudsman was established in 1962 [*Parliamentary Commission (Ombudsman) Act 1962,* consolidated and amended by *Ombudsman Act 1975*]. Part of the ombudsman's responsibilities is to investigate and review the decisions of Ministers of the Crown, government departments, and organizations covered by the Act and, also, decisions *not* to make available information as requested under the *Official Information Act*. The case notes issued by the ombudsman suggested that many of the cases investigated involved decisions in the area of discretion, where departmental rules and established precedents no longer have direct application and where public servants must exercise judgment. Here, *moral* considerations, such as respect for justice, equity and cultural differences, loom large. As a consequence there arose a perceived need for improved decision making by government officers, and the State Services Commission was requested to consider decision making in relation to the *Official Information Act*. For this purpose Robert Shaw was seconded to the Commission in November 1983 with a brief that by then had been extended to develop a project that would deal with decision making more generally, but at the higher levels of the Public Service.

The problem then was to devise a theoretically sound model of decision making that would address the real problems of public servants, and that could be incorporated into an effective training package. James Marshall and Michael Peters of the University of Auckland were contracted to advise upon the development of the model and the associated teaching/learning package and provide a substantive evaluation of the total project.

The project was undertaken between March and October, 1984, and comprised three main phases: an *initial* phase during which a project

team was established to help, in conjunction with the evaluators, to develop both the model and the teaching/learning package; an *experimental* phase designed to pilot the model and the teaching/learning package with a limited audience; and a *developmental* phase, consisting of six trials, designed to demonstrate the development of the skills of course participants.

The developmental trials involved 30 senior public servants representing 16 different government departments and major private sector interests (Fletcher Challenge Ltd.; Tasman Pulp and Paper Ltd.), in a total of six three-day workshops run over a period of three months between August and October 1984. The private sector was invited to participate to ensure that economic considerations were addressed explicitly and actively, and because of the potential wider application of the model of decision making.

The State Services Commission contracted with the evaluators to provide the following:

(1) assistance to the commission with the development of a model of administrative decision making;

(2) evaluation of the experimental trials in Auckland and Wellington and the production of advice on the development of the teaching package being produced, before the material for the developmental trials was finalized; and

(3) evaluation of the developmental trials in Auckland and Wellington.

This article is concerned with the evaluation of this project—its instruments, methodology, and major findings. It describes the evaluation of a *philosophical* (as opposed to a sociological or psychological) model of decision making and the associated teaching/learning package in the setting of the New Zealand Public Service, where a deliberate attempt has been initiated to improve the quality of decision making by government officers. In the next section we describe the model of decision making. Section III outlines the methodological approach and the formative evaluation. The summative evaluation and major findings are presented in section IV. A concluding discussion, section V, concerns itself with generalities learned from the project.

II. THE DECISION-MAKING MODEL

The model of decision making developed was initially based (Shaw, 1982) on an account of discretionary justice by Davis (1969), Wilson's

(1967) approaches to decision making as regards moral education, and Toulmin (1958) with regard to reasoning. Both the model and the teaching/learning package differ considerably from previous approaches in the area of decision making (e.g., decision or game theory). The model is based on a philosophical (as opposed to a psychological or sociological) approach to decision making. It addresses itself to the objective factors that should be examined in the *justification* of complex, problematic decisions, and does not *prescribe* an approved or ideal set of decisions but, rather, encourages the investigation of alternatives.

A major point of concern was not so much that the model eschewed psychological and sociological approaches to the *processes* of decision making but, rather, while embracing what is a logical approach to decision making, it ignored the considerable applications of theoretical or mathematical logic to decision-making theory. On this issue the arguments of Habermas (1971) are accepted.

Essentially, Habermas's arguments are directed against the technocratic notion of rationality that he claims underlies modern society, especially as exemplified in the gamut of systems, game, and decision theory. He claims that the growth of technology in social theory has led to the exclusion of normative elements from such discourse and that the actual technical procedures become established as values in themselves so that, for example, decision theory rationalizes the relationships between goals and values rather than simply being the technical means for reaching these goals. His criticisms of technocratic rationality, where goals are subsumed under specified means, are given in terms of four levels of increasing rationalization—the final level of which is when decision making, epitomized in game theory, is seen as a self-programming feedback system where wider value questions cannot be framed. Against decision theory the philosophical approach fo decision making in this model ensures that the normative and ethical elements are both recognized and open to critical, that is, rational, appraisal.

Furthermore, public servants are sometimes required to justify their decisions, that is, to give an account of the reasons that governed a particular decision. But, in theory these should be exactly the same reasons that were considered in the making of the decision—that is, at the same time of the taking of the decision the justification should be available. Even if sociological and psychological factors affect the *processes* of decision making the decision should still be justifiable according to whether the reasons adduced are indeed *sound* reasons. In the administrative-political world this kind of reason giving is appropriate and in practice accepted. Sociological or psychological accounts of how a decision was arrived at may be of interest to theorists but they cannot be used by administrators to establish that the "right" decision was reached.

It has been noted that talk of justification is not merely ex post facto justification (or worse, rationalization). Rather, it enters also into the making of the decision. When public officials are called upon to justify or review a decision already taken, the model provides a systematic method of ordering, recording, and investigating factors pertinent to the decision. A decision taken in accordance with the model must therefore provide a measure of accountability for decision makers that in itself will provide in part a justification for a decision should it later be called into question or reviewed.

It is timely that a project on the education of senior government officials attempts to focus on the ethical dimension of decision making. In the United States academics have written for more than two decades of the importance of placing ethics at the center of practical administration. Loosely called the "new age theorists," some wish to make radical changes in the way government organizations operate whereas others seek a more slight adjustment of the balance of competing concerns.

A strong case is made by Dvorin and Simons (1972) for placing human values and ethics at the center of practical public administration. They also wish to end the domination of theory by practice in bureaucracies, saying that operational machanics and political strategies should be replaced by a concept of the public interest, based on the individual as the most important concern of bureaucratic power. This implies that standard operating procedures must be replaced by more particularized procedures.

A similar conclusion may be derived from the work of Frederickson (1981). After summarizing the emergence of the "new public administration," he claims that one of its abiding features is the refreshing emphasis placed on values: administrators are not value-neutral, nor are administrative theories "value neutral models."

The model designed to address these issues is in part deduced from a concept of administrative discretionary justice (ADJ) derived from the work of Davis (1969), and which is defined as follows:

> Administrative discretionary justice is the resolution of the legal, moral, factual and other practical issues which arise in a situation where an official person acting in a capacity which is, in the final analysis, defined and limited by government legislation makes a just decision which results in action or inaction which affects other persons.

In this definition the notion of a "just decision" is minimally defined according to those precepts that define the notion of natural justice (Rawls, 1971). These are guidelines intended to preserve the integrity of the discretionary decision-making process. Thus a public administrative decision-making system must make provisions for conducting orderly investigations: It must have rules of evidence that guarantee rational procedures. The rule of law requires some form of due process—that is, a process reasonably designed to consider the interests of persons affected by the potential decision in ways compatible with the rule of law. For example, such a process would be independent and impartial, so that no person can judge in his or her own case, and the process must be fair and open and not prejudiced by pressure groups or public clamour. The precepts of natural justice are to ensure that the procedures of public administrative decision-making systems are impartially and regularly maintained.

With these definitions we deduced that an administrative decision made in accordance with discretionary justice would have considered these *types* of factors: legality, morality, facts of the case, just procedures, and possible decisions; to these can be added the implications or effects of decisions, and the strategies that might be used to implement a decision. Furthermore, it is assumed in these definitions that a decision has to be made and that the official is indeed the right person to make that decision. In other words, it is assumed that certain conditions have been fulfilled for the model to even apply.

The components of the model were then renamed slightly and were defined as follows (in the version of the model to emerge at the end of the project):

(1) *Conditions:* those criteria that must be satisfied if the model is to be applied. The conditions do not in any way guide a decision; they state when it is proper for an official to make a decision.

(2) *Facts of the case:* the circumstances, constraints, relevant personal information, and the likely consequences of particular decisions, that when taken together describe a particular case. Consideration of veracity of facts, and the absence of information falls within this component.

(3) *Rules:* the rules relevant to, or that "govern" a decision, which may be found in a constitution, in legislation, bylaws, accredited manuals, or informally drafted written rules, decisions, and precedents.

However, even when complete, formalized, carefully drafted rules are available they do not determine the decision outcome, as discretion still

applies. Formal rules may, however, indicate that certain answers are wrong. If the rules discrediting a decision are found in law, the answer is wrong in law. The determination of which rules (particularly precedents) will be counted as relevant in a particular case may sometimes give rise to complex conceptual problems that may need to be interpreted in terms of the departmental ethics.

Note here the extension of the term "discretion" to areas where decision outcomes should be prescribed by existent rules, in order to cover those instances where (say) legislation is outdated or about to be replaced, or is in need of replacement (see Davis, 1969). This is compatible with the tenets of "the new public administration."

(4) *Morality:* first, the moral principles held by the organization or the decision maker to be potentially relevant in any application of ADJ, and second, the persons or parties affected (together with an account of how they are affected) by a potential decision.

(5) *Procedures:* guidelines intended to preserve the integrity of the discretionary decision-making process. A public administrative decision-making system must make provisions for conducting orderly investigations; it must have rules of evidence that guarantee rational procedures. The rule of law requires some form of due process, that is, a process reasonably designed to consider the interests of persons affected by the potential decision in ways compatible with the rule of law. For example, such a process would be independent and impartial, so that no person can judge in his or her own case, and the process must be fair and open and not prejudiced by pressure groups or public clamour. The procedures, or the *percepts of natural justice,* are to ensure that the procedures of public administrative decision-making systems are impartially and regularly maintained.

(6) *Options:* all of the alternative decision outcomes available. It is in this component that the major conclusions are reached.

(7) *Implications:* those matters that may affect the implementation of the decision once it is made, and which fall outside the control of the decision makers once the decision is made. The implications cannot, by definition, be completely assessed before the decision is made. They may sometimes be important in explaining what went wrong.

(8) *Strategies:* practical proposals developed to assist with the successful implementation of the decision proposed.

With each logical component, key checklist questions for consideration have been devised. For example, within the morality component the key questions are as follows:

(i) What are the principles of morality that should be consciously considered by officials when the decision is made?

(ii) Who are the parties affected by potential decisions?

Further questions under each of these key checklist questions have also been developed.

It can be seen that if an official follows the "sequence" given by the model and the checklist questions, then there will be a systematic ordering of relevant "information," an analysis of the problem in terms of the model, an indication of any extra "information" to be sought as a result of reminders from the checklist questions, and, it is hoped, the emergence of possible options.

However, although essentially a logical model, there is no claim that the model is an algorithm for calculating or reading off the correct decision. The traditional logical problem of deriving an "ought" from an "is" is but one daunting prospect underlying that possibility, and even if normative and prescriptive propositions were fed into the processes of any argument designed to permit the deduction of a decision, arguments on normative preference would still abound over the adoption of substantive normative and prescriptive premises. The assumptions underlying the model then were that it would not provide an easy formula for decision making. On the other hand, the model, employed in Popperian fashion (Popper, 1963), provides principles of criticism that when employed in a critical manner can throw out wrong decisions. For a fuller discussion of the model and its justification see Marshall et al. (1985).

III. METHODOLOGY

RATIONALE

The term "evaluation research" covers a wide range of different methods and approaches from classical experimental design to the more recent social-impact assessment and as such does not represent anything like a methodological orthodoxy (Struening and Guttentag, 1975). In general, however, evaluation research may be defined simply as

> the systematic application of social research procedures in assessing the conceptualization and design, implementation and utility of social intervention programs [Rossi and Freeman, 1982: 20].

Given the practical nature of the project and the involvement of various governmental groups in its development and implementation, it

was considered essential that, in general terms, the methodology to be adopted should be:

(1) *developmental,* where each step is partially determined by the knowledge gained in previous stages in accordance with broad, long-term project objectives;

(2) *collaborative,* seeking the cooperation and ensuring the possibility of participation of all those involved in the project;

(3) *practically oriented,* where evaluation data gained in a series of iterative feedback sessions is directed at framing both a model of administrative decision-making and an associated teaching/learning package; and

(4) *flexible,* allowing for the adoption of a wide range of appropriate social research techniques.

An evaluation of this sort requires a case study approach focusing on questionnaire and interview techniques in order to acquire a combination of qualitative and quantitative data.

Finally, this model was associated with a teaching/learning package designed to meet educational criteria that

(1) although theoretical, the model be grounded in practice;

(2) it assist government officers to analyze decisions in a systematic manner;

(3) it be based upon case studies of actual decisions made; and,

(4) it should not teach senior government officers "to suck eggs" as they would claim, in virtue of their seniority, to be experienced decision makers.

How the package met these criteria will be discussed in the following sections. For a fuller discussion of the educational problems and the "resolution" see Marshall and Peters (1986).

The project clearly distinguished between the need for both a *formative* and *summative* evaluation. The formative evaluation, which required the evaluators to analyze project processes and outcomes with a view to improving the theoretical aspects of the model, the effectiveness of the teaching/learning package, and the possibility of achieving its objectives *during* the life of the project, called for a consultative Action Research Approach (Peters and Robinson, 1984).

The summative evaluation, it was thought at the beginning of the project, would develop out of the accrued and combined experience of project team meetings and experimental trials. In general, it would demand the development of a specific evaluation process and specific evaluation instruments not only to indicate pre-/post-course changes of participants' attitudes, but also to demonstrate some degree of internal

validity by providing a series of cross-checks of the perceptions of all groups involved—the evaluators, the project team, the course participants, and the course director.

The evaluators, on negotiating the contract, submitted a research proposal outlining the above methodology and specifying key research questions in accordance with the planned major phases of the project. In addition, they identified the need for an initial evaluability assessment. The evaluability assessment, in general, is designed to establish a favorable climate for the evaluation by identifying what, if any, difficulties stand in the way of carrying out a successful evaluation (Wholey, 1977). Such an assessment not only provides a general orientation to the evaluators but also serves to clarify objectives, establish initial agreement and collaboration with those involved in the project, and clarify criteria of the evaluation and the priorities of investigation.

THE EVALUATION PROCESS AND INSTRUMENTS

Both the formative and summative evaluations are best conceptualized in terms of a series of iterative feedback cycles involving the evaluators, course directors, project team members, and course participants, according to the Action Research Model (see Peters and Robinson, 1984).

The evaluation process can be best appreciated in terms of Figure 1. The model incorporates different levels of evaluation, symbolized by roman numerals. Each level is characterized by different forms.

Level I

(i) Background information: a form recording relevant personal information, qualifications, and work experience of individual course participants.

(ii) Decision-making skills questionnaire: a questionnaire (given as pre-and posttest) based on a seven-point scale developed from components and questions of the model of administrative decision making.

(iii) The immigration case: this case is introduced during the half-day preparation for the three-day teaching workshop for two purposes: (a) as an exemplar of a case-study to help course participants prepare and develop their own case-study material; and (b) for evaluation purposes. The evaluation is structured around the course participants' responses to each of the four decision points. Participants are asked to list factors that they considered in reaching their decision, to weigh

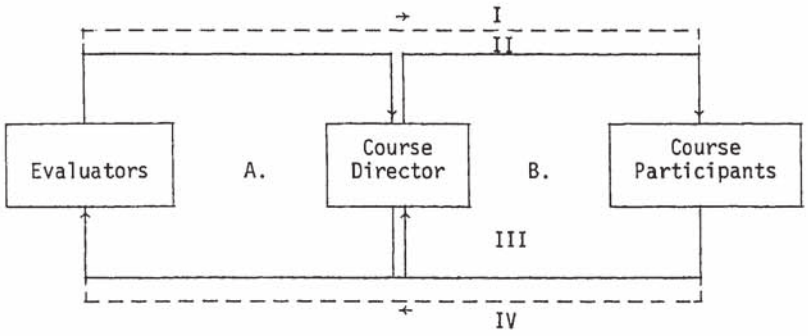


Figure 1: Model of the Evaluation Process

factors according to a seven-point scale (according to importance), and to write a draft letter as the decision maker. A special form is provided. Each of the forms of evaluation at this level were developed by the evaluators, and submitted to project team members for comments (cycle A). They were to be administered by course directors.

Level II

(i) Course director's report: This evaluation takes the form of an open-ended progress report written at the end of each day. It provides an opportunity for the course director to comment on the progress of the course participants (considered as a group) toward an understanding of the model, and to make any further observations that he or she might consider pertinent. A special form is provided that lists criteria course directors might consider in writing up their reports.

Level III

(i) Feedback meeting: this is a taped two-hour meeting held at the end of the three-day developmental trial (chaired by the course director), which is designed to elicit comment and general response from course participants on: (a) the model—its statement definition, use, applicability, and so on and (b) the teaching/learning package—its effectiveness, strategies, structure, organization, and so forth. A list of potential questions for course directors is provided.

(ii) Test of knowledge of components of the model: the test, based on items of the questionnaire that itself was drawn from components of the model, simply instructs participants to classify items according to the components.

Level IV

(i) Final evaluation instrument and post-course structured interviews: The final evaluation instrument is framed in terms of a series of limited options designed to tap participants' perceptions on the usefulness of the course and their own implementation of the model. The selected interviews held with course participants one to two months after their attendance at the course were designed to further investigate the "transfer" of decision-making skills to the work environment. The interviews were designed to elicit open-ended responses to a series of structured questions.

These instruments were designed to give overlapping and cross-checking assessments of the perceptions of those involved—a feature required by the underlying collaborative nature of the evaluation. Further, the instruments were designed so that future courses can be self-evaluating.

IV. THE FORMATIVE EVALUATION

The model and the associated teaching/learning package were developed by a project team consisting of the project director, the evaluators, and staff training officers and line managers drawn from a variety of government departments. They met first on 2 March 1984, and thereafter twelve times during the ensuing months. The evaluators attended all but two meetings and, in addition, held independent meetings with the project director.

Part way through the project some members of the team withdrew and the responsibility for the development of the model fell more heavily on the evaluators and the project director. Some members had expressed doubts about their role and about the theoretical basis of the model (e.g., it did not address psychological or sociological issues). In the view of the evaluators the former doubts were justified to some degree as no formal role for the project team was established, and this may, in part, have led to subsequent confusion; the latter doubts arose from a lack of understanding of the model and its assumptions.

The initial version of the model was based on a research paper examining the growing concern in the academic literature over the role of ethics and values in public administration (Shaw, 1982). The model originally possessed a number of conceptual categories to be consulted in a systematic approach to the justification of complex decisions. The

evaluators commented on the original paper, suggesting the following: a fuller analysis of "discretionary justice"; retaining the conceptual categories but checking for conceptual overlap and "tightening up" of definitions; and developing key and checklist questions under each category. Further, they suggested that a Popperian (Popper, 1959, 1963) variation might be developed that would stress the essential fallibility of decision making and its improvement through trial and error, and active challenging criticism designed to overthrow potential tentative decisions.

These suggestions were later formatized in the evaluability assessment submitted on 16 May along with a full definition of the central concept of discretionary justice and an operational version of the model fleshing out the earlier suggestions made in the original paper. In the evaluability assessment the evaluators did the following:

(1) identified theoretical problems in the model (mentioned above);

(2) noted effects of the delay in producing the teaching/learning package upon the evaluation, especially in developing and piloting evaluation instruments in the experimental trials; and

(3) noted problems for the effectiveness of the summative evaluation.

As regards the last point, it was argued that the planned three-day workshops for the teaching of the model were of such short duration that it was unlikely that there would be any significant behavioral changes, or changes in skills, that could be picked up on pre/posttesting. Further, the evaluators asserted that it was not at all clear at that stage of the proceedings that the model could be translated into a body of *identifiable* and quantifiable skills. They added, however, that it *should* be possible to assess changes in attitudes toward decision making and attitudes toward the model.

The development and structure of the teaching/learning package emerged in a usable form from the experience of the experimental trials. A half-day introduction consisting of an introduction to the model by use of the immigration case—a case drawn from the ombudsman's case notes—which provided participants with an exemplar of a prepared case study and an indication of how administrative decisions could, in the context of a course, be examined.

After the experimental trials, the immigration case—a case involving the disputed levying of duty and sales tax on the imported household goods of a Chinese couple immigrating to New Zealand—was given a

standarized presentation in order to assist prospective course directors and to realize the demands of evaluation (see section III).

The structure of the three-day workshop, following the half-day introduction, consisted of the presentation of case studies by participants (who had prepared them from their departmental files in accordance with a standarized format) to their colleagues, while the course director used this material to highlight aspects of the model and inductively work toward analyses of each decision presented in terms of the components of the model.

Each case study was structured in terms of a series of decision points and participants were encouraged not only to formulate their own decisions at each point but to debate them with their colleagues and then to reflect further upon their individual decisions before the actual decision taken was revealed. Following was a critical discussion of the actual decision taken, and the structure was repeated for the next decision point. Much, therefore, depended upon the inherent interest of the case studies selected and their adequate selection, preparation, and presentation. On the second day of the workshop there was a formal teaching session on the model followed by further case studies, which were subjected to increasingly rigorous analysis and critical discussion. Role-playing between participants was used with varying degrees of success. A video presentation of the model was developed during the project.

On reflection and from feedback two crucial issues emerged: first, the timing and sequence of case studies; and, second, the required judicious *balance* of the presentation of individual case studies and the formal teaching session. Both these matters called for mature judgment by course directors and highlighted the need for adequate training of course directors.

Participants in the trials gave the evaluators positive feedback about the model and the structure of the course. The trials proved to be a vital and necessary stage in the development of both the model and the teaching/learning package. Specifically, they provided the opportunity for the project team (including the evaluators) to: develop and trial the teaching/learning package; observe and participate in the teaching of the courses; trial, develop further, and gauge response to the model of decision making; and to formalize the staging, sequence, and presentation of case studies. Further, as a result of reflection on the experimental trials, the evaluators were able to develop a standardized format for the immigration case and to develop specific evaluation instruments.

During the formative evaluation full running records were kept of all meetings. These were recorded in the form of summaries made from official minutes, taped sessions, and the evaluators' notes. The summaries detailed major phases of the project; dates and types of meetings; participants involved; all documents referred to or used; and major decisions and outcomes. They were submitted to the project team for verification. In addition, the evaluators wrote a series of formal letters to the project director concerning various major issues of the evaluation. Project team members reported favorably on the value of these communications.

V. THE SUMMATIVE EVALUATION: MAJOR FINDINGS

Certain problems were experienced with the administration of the instruments. Some officers were reluctant to participate in the evaluation. Perhaps the point that it was *the course and its materials* that was under evaluation was not stressed firmly enough. Finally, the success of the administration of the instruments was determined by the structuring of the course by the course director—in some cases pedagogic concerns overrode the interests of evaluation. We restrict ourselves in what follows to a summary of the major findings of the summative evaluation presented in the order given in the model of the evaluation process (Figure 1). The *level* of evaluation is indicated by roman numerals, and the *form*—corresponding to individual evaluation instruments—by small, bracketed, roman numerals (see methodology section).

I(i). The Background Information Form revealed certain salient characteristics of government officials participating in the trials. The average age of the participants was 42.4 years (ranging from 29 to 59), and the average number of years of those in the public service was 21.8 years (ranging from 6 to 38). Most of the course participants indicated that they had experience in decision making in the areas concerning finance/accounting and staffing/personnel. Recorded expectations revealed that most participants were unclear as to the nature of decision making involved in the course and appeared to expect a course on psychological/sociological processes.

I(ii). The questionnaire was an attempt to gauge changes in attitudes toward factors associated with the exercise of discretion in decision making. For this reason the analysis of the questionnaire was directed toward participants' perceptions of moral factors, the exercise of discretion and responsibility, and the importance of cultural differences. Items for the questionnaire were drawn from each of the components of

the decision-making model. They were drawn specifically from the checklist questions developed under each of the ten components, and arranged in two groups of 20 items covering decision-making abilities and attitudes (see Appendix 2).

Participants were asked to rate each item on a 7-point scale for their importance in general and for their present position. The questionnaires were administered to participants before and after the three-day trial (pre- and post-course). Some rogue items were included. The questionnaire results were recorded for each trial participant indicating individual shifts in the ratings of two sets of 20 items. Individual scores were averaged for each trial and pre/post-course rankings were established for decision-making abilities and attitudes.

Participants *consistently ranked moral factors very low* (almost always in the last ten items and often in the last five). This finding was further confirmed independently in the analysis of selected items.

Some improvement in the ranking of moral factors, however, was evident after the development trials. For the trials as a whole, moral factors in the abilities category improved up to *two* placings. Although in general it can be said that the developmental trials had some positive effect in bringing about an awareness of moral factors in the area of decision making involving some element of discretion, it would be difficult to make inferences from this data to post-course changes in working behavior.

The small degree of improvement in ranking of moral factors must be set against other considerations, such as the relatively short duration of the trials, the effectiveness of the teaching package, and the participants' knowledge of the model.

Improvements in the rankings of items concerned with the exercise of discretion appeared to be at the expense of rule-following items.

I(iii). Participants were asked to list all factors taken into consideration at each decision point in the immigration case and to rank each factor according to a seven-point scale. Further, participants were asked to draft a response in the form of a letter to what prompted the decision at each point. The analysis of the results are too detailed to be included here (but are available on request). Only two participants clearly wished to overturn the decision at the first decision point on grounds of cultural difference (one major ground on which the ombudsman eventually brought down a finding against the department concerned). By decision-point three, acting as ombudsman, most participants wished to review the original decision, though on grounds of departmental administrative inefficiency (delays in correspondence, lost files, etc.) and doubts concerning the onus of proof on proving whether household goods were new or used. The results of the immigration case offered some confirmation of the low ranking of factors of cultural difference by participants.

II(i). The perceptions of the course directors were seen as important for the collaborative aspect of the evaluation and for the purpose of providing a cross-check on the perceptions of the evaluators and the participants. They also provided an opportunity for the course director to reflect upon the proceedings of the day, and to isolate problems in the teaching of the model and the presentation of case studies so as to restructure the next stage. The full potential of this instrument was not realized, although course directors' reports confirmed the evaluators' final recommendations on the teaching/learning package.

III(i). Feedback meetings for each trial were conducted at the end of the three-day workshop. Generally, participants were favorably disposed toward the course.

Participants stated that the model structured and systematized their "unconscious" processes of decision making; it captured their experience as public administrators. Yet all participants had trouble with some of the components, including problems with titles, definitions, and conceptual overlap between components. They confirmed the model was best suited to complex, problematic decisions and thought it most useful and applicable at top and middle levels of management within the public service. The major objections to the model were based upon problems of time and efficiency in applying the model.

Many participants commented that they saw the interdepartmental interaction as crucial for the course as it encouraged critical discussion and promoted the group dynamics.

Finally, some concern was expressed over the post-course implementation of the course—a matter that was investigated by the evaluators in the post-course structured interviews.

III(ii). Test of the knowledge of components of the model was as follows: Participants were asked to identify the appropriate component of the model to which items on the Decision-Making Skills Questionnaire might be allocated. In analyzing this data the evaluators took account of possible overlapping components and discounted some items as rogues. Only one candidate scored considerably less than 50% and no one candidate scored higher than 75%. The average was 53%. Test scores related well to course director assessments and a score of 53% was not regarded as a satisfactory grasp of the model.

In a similar test administered in the post-course structured interview—admittedly to a small number of candidates—there appeared to be a slight drop only in participants' scores on knowledge of the model.

The analysis of test results indicated those components of the model least well known (*Procedures, Options* and *Strategies*) and those known best (*Morality*). This is an important result when compared with other data collected for it indicated that participants seemed both to know *moral* components well *and* to rank them low relative to other types of factors.

Finally, it was discovered (Final Evaluation Instrument) that participants' perceptions of their understanding of the model was actually higher than their performance.

IV(i). The Final Evaluation Instrument was designed, along with post-course structured interviews, as "follow-up" assessments. Participants were asked to rank various items covering the usefulness, implementation, and understanding of the model on a seven-point scale. It was administered to approximately one-third of all participants.

In general, participants were very positive toward the course. Two items—the course as both providing an introduction to a systematic approach to decision making, and an opportunity to work on case studies from other departments—were consistently ranked very highly.

V. CONCLUDING DISCUSSION

The final report including a nontechnical summary, conclusions, recommendations, 23 detailed appendices, and a set of evaluation instruments for the future assessment of the teaching/learning package was submitted to State Services Commission in December 1984. The Commission distributed copies of the report to selected New Zealand academic libraries.

The evaluators made a number of recommendations concerning both the model and the teaching/learning package that detailed the changes, which in their investigations, were deemed necessary for the smooth operation and organization of future courses. The Commission in turn, has decided to run a series of 10 courses in 1985 in administrative decision making for both senior executives and middle managers in the public service based on the teaching/learning package developed and a further 12 courses in 1986. In addition, given the deliberate low ranking of moral factors vis-à-vis other factors in decision making, the evaluators strongly recommended that the Commission undertake further research in this area.

Given the nature of the project, and the methodology adopted, what general learnings emerged from this evaluation? This is an especially

difficult question, as in attempting to specify learnings that transcend a particular project often all one is left with is a set of methodological guidelines that prove to be so general as to be almost empty. Yet because this field is so young compared to other fields of social science research and because it is fast becoming an indispensable tool, not only in the development of staff training programs but in the formulation, more generally, of social policy, it is worthwhile to reflect on general matters.

We approach this question simply in terms of a number of listed problems and statements:

(1) The collaborative-consultative action research approach provides a suitable perspective for evaluation projects of this kind, for it allows the necessary flexibility in the adoption of research techniques and captures the essential developmental nature of projects evolving through time. This feature is of primary importance. It describes the process of evaluation as a series of iterative feedback cycles where the evaluators, in dialogue with others, can trial or pilot various ideas and learn from mistakes. It is not the sort of methodology that prescribes a set of strategies that must be followed to the letter. Action research is action-oriented and problem-focused. As such it provides no magical methodological solutions, but is proof to all of the real-life complexities that occur when people collaborate over a period of time in order to develop, organize, and put into practice a series of general, long-term objectives only perceived in outline at the beginning.

In the present project, and probably in most similar projects today, it is not difficult to gain agreement that course structures and evaluation methodologies should mutually evolve—but it was found that, in practice, project team members and course participants were not so willing to allow that the *objectives* themselves should evolve. The evaluators acknowledge that team members may feel "lost" when asked to contribute in determining the very purposes of the project, but also affirm that this is a legitimate and productive way of proceeding.

(2) Although the action research approach provides a general methodology (and an implicit developmental learning theory) it does not obviate the need for the recognition, design, and development of specific evaluation instruments required by particular projects. Such instruments, given the developmental approach, can be piloted, tested, and modified in accordance with both feedback from the project team and the experimental results.

(3) A set of problems experienced in this evaluation concerned the way in which project team members perceived their roles. It is important that all groups of participants know initially, if only in outline, what is expected of them, especially when there are more than two groups involved. Although the formal specification of roles may, in part, overcome any role confusion, it is important to realize that such role specification needs continuous examination and consideration according to changes in the conception and direction of the project as it evolves.

(4) In conjunction with the stated problem above there exists the difficulty of developing and maintaining effective lines of communication between the various groups involved in the evaluation. Not only does this imply attention to group

processes but also the recognition of the official status and position of individuals within an organization. Those people being introduced to a novel or "foreign" way of thinking may have a presumed expertise in the area. Unless the evaluator understands something of the way participants perceive themselves, the evaluator runs the risk of causing an initial hostility or threat that may be difficult to overcome. Further, unstated political biases or assumptions, and thus the potential for friction between various factions, may work against the interests of the project and the achievement of the project's objectives if the evaluator and the participants fail to perceive that much misunderstanding results from a failure to recognize or clarify ideological stances or implicit assumptions.

The formal documentation of all communications is essential. Such documentation as, for example, official minutes of meetings and correspondence are of paramount importance in the reconstruction of events and outcomes. Often the evaluator may have to officially elicit a response to a suggestion or recommendation in order to determine who has the responsibility for making the next move.

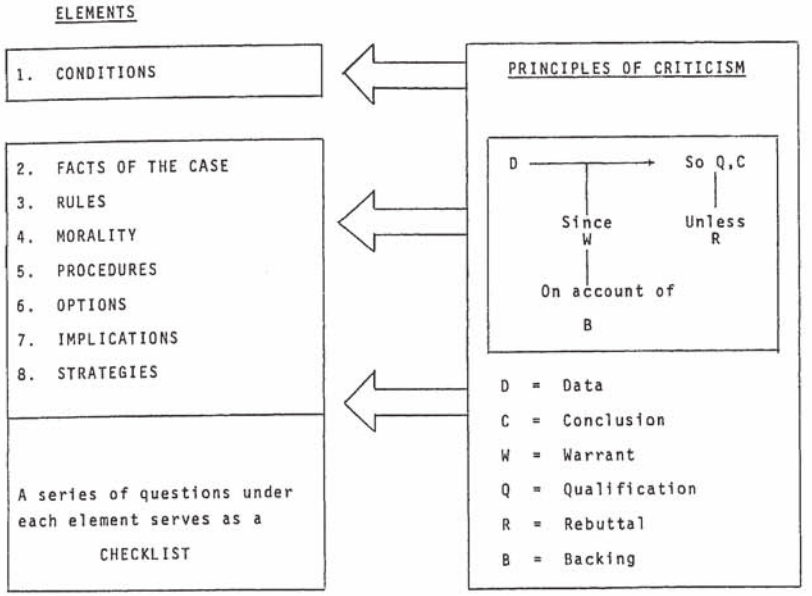
(5) In this evaluation the question of just what a body or set of cognitive skills might consist of was raised a number of times. It is an important problem that must be addressed if genuine progress is to be made in the training of senior public servants. This is not the place to consider the psychology of learning, nor is it the place to discuss decision-making skills as such, but the evaluation of the success of decision-making courses should be undertaken in forms of concrete observable criteria of behavior change and these will only be developed as the result of a conceptual work undertaken in the light of observations.

(6) The problem of on-the-job implementation of learning from courses is a general problem evidenced throughout the literature. This project had a strong orientation toward the practical and the relevant, employing as it did a case study approach based on actual (and often ongoing) decisions. In trials some participants (most frequently those from district offices) revised their views of actions. There was one person who rushed off to check all his financial delegations in the middle of a course! The learnings that were in and of the model were often subtle and not easily evinced by test. As has been indicated, the success of the project in this area has been limited, but the evaluators and the Commission personnel involved remain committed to the importance of assessing such learnings.

(7) Finally, we wish to comment on the close conceptual interrelationship of two sets of questions in this project: an educational set and a set concerning evaluation. The first set involves issues essential to the evaluation and to the eventual recipients of both the model and the teaching/learning package: For example, how should we "educate" senior officials in an area in which they would claim to have considerable experience? In the second set, were traditional issues associated with any evaluation? Given the bewildering number of possible approaches to evaluation, what methodology should be employed? As the project developed the two sets of questions became indistinguishable; at almost every practical turn evaluation questions became educational questions, and vice versa. On reflection, educational questions were seen to be an inherent part of, and central to, the evaluation because the project director and the project team who were to develop the model in conjunction with the evaluators, in effect, constituted a *learning community*. Educational questions defined the ground and form that the evaluation was to take.

On further reflection we began to speculate about a form of evaluation that took educational problems seriously, to the extent that they should form the theoretical framework of any evaluation (see Marshall and Peters, 1985).

APPENDIX 1
The Logical Model



SOURCE: Toulmin (1958). Reprinted by permission of Cambridge University Press.

APPENDIX 2
Decision-Making Questionnaire Items

Abilities	*Attitudes*
Ability to:	Willingness to:
(1) identify the need for a decision	(21) identify approprate department
(2) delegate a decision	(22) delegate a decision
(3) consult with others	(23) consult with others
(4) distinguish relevant facts	(24) accept responsibility for decision
(5) gauge likely consequences of decision	(25) gauge likely consequences of decision
(6) weigh the evidence	(26) seek additional information
(7) view information impartially	(27) view information impartially
(8) identify precedents	(28) determine relevance of legislation
(9) interpret departmental policy	(29) exercise discretion
(10) exercise discretion	(30) make a quick decision
(11) make a quick decision	(31) distinguish different interests
(12) distinguish different interests	(32) advocate departmental policy
(13) identify a conflict of interests	(33) action decision
(14) identify moral percepts involved	(34) admit mistakes
(15) question previous decisions	(35) reach a compromise
(16) action decisions	(36) record all pertinent matters
(17) forsee limitations	(37) follow set procedures
(18) reach a compromise	(38) review decision
(19) follow set procedures	(39) take account of cultural difference
(20) exercise moral choice	(40) exercise moral choice

REFERENCES

DAVIS, K. C. (1969) Discretionary Justice: A Preliminary Inquiry. Baton Rouge: Louisiana State University Press.

DVORIN, E. P. and R. H. SIMONS (1972) From Amoral to Human Bureaucracy. New York: Harper and Row.

FREDERICKSON, H. (1981) New Public Administration. University of Alabama Press.

HABERMAS, J. (1971) Knowledge and Human Interests. Boston: Beacon.

MARSHALL, J. D. and M. PETERS (1986) "Evaluation and education: practical problems and theoretical prospects." New Zealand Journal of Educational Studies.

———(forthcoming) "Evaluation and education: the ideal learning community." Policy Sciences.

MARSHALL, J. D., M. PETERS, and R. SHAW (1985) "Administrative discretionary justice: a model of decision-making for the public service." Unpublished manuscript available from the authors.

PETERS, M. and V. ROBINSON (1984) "The status and origins of action research." J. of Applied Behavioral Sci. 20: 113-124.

POPPER, K. R. (1959) The Logic of Scientific Discovery. London: Hutchinson.

———(1963) Conjectures and Refutations. London: Routledge & Kegan Paul.

RAWLS, J. (1971) Theory of Justice. Cambridge, MA: BEL-KNAP Press.

ROSSI, H. and H. E. FREEMAN (1982) Evaluation: A Systematic Approach. Beverly Hills, CA: Sage.

SHAW, R. S. (1982) "Administrative discretionary justice." Presented to the annual conference of the New Zealand Association for Research in Education.

STRUENING, E. L. and M. GUTTENTAG [Eds.] (1975) Handbook of Evaluation Research, vol. 1. Beverley Hills, CA: Sage.

TOULMIN, S. (1958) The Uses of Argument. Cambridge: Cambridge University Press.

WHOLEY, J. S. (1977) "Evaluability assessment," pp. 41-56 in L. Rutman (Ed.) Evaluation Research Methods: A Basic Guide. Beverley Hills, CA: Sage.

WILSON, J., B. SUGARMAN and N. WILLIAMS (1967) Moral Education. Harmondsworth, England: Pelican.

James Marshall is Senior Lecturer in Education at the University of Auckland. He received his Ph.D. in Philosophy from Bristol University in 1973. Dr. Marshall is currently the editor of Access, *and has published widely in education and policy journals.*

Michael Peters is Honorary Research Fellow and Temporary Lecturer in Education at the University of Auckland. He received his Ph.D. in Education from the University of Auckland in 1984, and has published in education and policy journals.

Robert Shaw is Education Officer in the Examinations and Assessment Units of the New Zealand Department of Education in Wellington. He received his Ph. D. in Education from the University of Auckland in 1976, and has a master's degree in Public Policy from Victoria University of Wellington. Dr. Shaw has published in the Journal of Moral Education *as well as other New Zealand journals.*