

Position Paper: Program Evaluation and Quality Control

Ronald Mackay

It is probably true to say that the topic of techniques for carrying out evaluations of language programs has not received as much attention as have the other topics dealt with in these EXPLORATIONS Seminars (with the possible exception of the first in this series, which examines the role of research in the formulation of language policy).

There are many reasons why the topic of program evaluation has received less attention than, say, program design or teacher training or classroom techniques. Evaluation can be regarded as of a lower priority than the more obviously immediate activities associated with design and planning. Evaluation can be regarded as something to be done in the future, once the program is 'ready' for it. And with budget priorities being given to the more obviously creative aspects of program design, and the constantly evolving nature of the planning function, that future never comes and the program is never quite 'ready' for the scrutiny of the evaluation step.

Another reason for neglect is to be found in a common perception of the nature of evaluation. In this commonly held perception, program evaluation requires those who have been most intimately involved in its conception, design and implementation to step aside and abandon their creation to the attention of an alien, outside individual or group of individuals who will measure it, analyze it, judge it, and present their findings in a report replete with statistical tables not easily understood by the program stakeholders for whose enlightenment the evaluation was purportedly carried out.

This is an unfortunate and inaccurate perception of program evaluation. Unfortunate, because it encourages those responsible for program design and planning to hold evaluation at arms length instead of embracing it as an important and valuable management aid. Inaccurate, because it does not do justice to evaluation as a cooperative venture involving the active participation of program personnel and the evaluation specialist in an endeavour designed to answer questions of importance to program stakeholders.

It would be counter-productive if not quite misleading to attempt to provide a unique formulaic set of procedures for program evaluation.

There is no single recipe or predetermined blueprint for the evaluation of language programs. This is not to say, however, that evaluation is an ad hoc, unprincipled, or arbitrary activity, dependent for its design or process on the whim of those involved in the conduct of the evaluation. I believe, however, that there is a basic logic underpinning program evaluation and that a familiarity and understanding of that logic on the part of all those concerned is important if any particular evaluation study is to be a principled activity as opposed to an arbitrary series of decisions and steps and activities.

It is, therefore, appropriate at this point to discuss the nature of the logic that underlies program evaluation at least as I see it, so that the principles are clear and can act as a guide to design in any particular case.

I have found the most appropriate way of presenting the principles underlying program evaluation to be to address a series of basic questions which are invariably asked, in my experience, by program personnel in the early stages of planning an evaluation.

WHAT IS MEANT BY “PROGRAM EVALUATION”?

Program evaluation is an activity involving:

- the purposeful and systematic collection, analysis and interpretation of information, about one or more components of a particular program, in order to address the practical concerns and answer the questions of one or more specific individuals or groups so that these answers can be used by the individuals or groups concerned, for one or more of a variety of predetermined purposes.

WHAT IS INCLUDED IN THE TERM “LANGUAGE PROGRAM”?

The creation and implementation of a second language program involves a complex combination of:

- decisions and plans: e.g. at what grade level language teaching will begin; who will teach it, specialist or generalist; how much time will be devoted to it in the school timetable, etc.
- resources: e.g. money, space, time and people available for the projects to be carried out
- activities: e.g. establishing committees (consultative committees, planning committees, production committees); distributing drafts of the program to teachers, etc.; obtaining feedback; conducting needs analyses
- people: e.g. politicians; administrators; curriculum design experts; consultants; teachers; students; parents; members of the community; materials writers; textbook authors; publishers; teacher-training

institutions

- interactions: e.g. between the politicians and the larger community; consultative committee and the production committee; authors/publishers and the funding body; teachers and students
- outcomes: e.g. student performance on teacher-made tests/ regional or provincial exams; teacher satisfaction with the approach/the materials; student satisfaction with the instruction/the materials.

WHAT COMPONENTS OF A PROGRAM CAN BE EVALUATED?

Any and all of the various stages or components of a language program can be evaluated. The stage or components that should or could be evaluated in any particular evaluation would depend upon the information requirements of the specific individuals or groups who requested that the evaluation be carried out.

These groups or individuals are usually called **STAKEHOLDERS** (i.e. persons having a 'Stake' or interest in the outcome of the evaluation). The persons or groups who most need the evaluation (i.e. those who have the authority or decision-making power to act on its findings) are usually called the **PRINCIPAL STAKEHOLDERS**. The principal stakeholder is usually (but not always) the program manager, that is, the person who has superordinate responsibility for the creation of the program and its implementation.

WHO ARE THE INDIVIDUALS OR GROUPS (THE STAKEHOLDERS) FOR WHOM AN EVALUATION OF A LANGUAGE PROGRAM MIGHT BE CARRIED OUT?

Potentially, an evaluation of a program can be carried out for any stakeholder or group of stakeholders, directly or indirectly affected by the program and by the results of the evaluation, e.g. politicians, the funding body, the program manager, administrators, teachers, principals, teacher-training institutions, parents, students, materials writers/authors and publishers.

In practice, evaluations are usually carried out for stakeholders who commission them, pay for them and require the evaluation results in order to make specific decisions about certain courses of action.

WHAT MIGHT BE THE PURPOSES ADDRESSED BY A PARTICULAR EVALUATION?

The purposes of a particular evaluation relate directly to the questions that a specific individual or group (i.e. the stakeholder(s)) wants answered. Given the variety of potential principal stakeholders and the

variety of interests or concerns any one of these might have, many different questions are possible, e.g.:

- a teachers' association might commission an evaluation to find out if its members are comfortable using a new program and if they are not, to provide it with a justification for additional in-service training
- the funding body might commission an evaluation to allow the program manager to find out how teachers and students are reacting to a new program so that he/she can decide where improvements or modifications are needed
- a Parents' Association might commission an evaluation of an innovative program so that they can decide whether to leave their children in it or return them to the regular program
- a publisher might commission an evaluation of a program in which the materials it publishes are being used to find out whether they should be revised and if so, in what ways.

WHICH OF THESE STAKEHOLDERS' INTERESTS HAS PRIORITY IN A SPECIFIC EVALUATION?

Normally, priority is given to the interests of the individual or group who has the authority to decide what questions the evaluator will address and to make recommendations or to make decisions based on the results of the evaluation. This is usually the individual or group who commissions and pays for the evaluation to be carried out.

However, it is also possible and highly productive to form a "working group" made up of all or some of the stakeholders. Together, the stakeholders can decide upon common concerns which could usefully be addressed by the evaluation. Alternatively, the working group of stakeholders can agree upon the order of priority of the questions that the evaluation should answer.

There are advantages to be gained by involving a range of stakeholders in the matter of determining the focus of the evaluation. The principal advantage is the increased participation of the stakeholders in an evaluation whose focus reflects their interests.

WHEN IS IT APPROPRIATE TO EVALUATE?

Evaluation — or evaluative activity — can be carried out at any time depending upon the component of the program about which information is being sought and how that information is to be used. For example, program goals may more appropriately be evaluated immediately they have been formulated and before the pedagogic materials have been developed. If changes are required in the goals, they can be made at a time

when it is easiest and least expensive to make them.

Likewise, the evaluation of the design of teaching units and exercise or activity types may be most effective if carried out as the prototypes are being designed. Immediate feedback can be used to adjust the design or the activities to provide an improved pattern for future units.

The effectiveness of the program as a whole, however may be most effectively determined after it has been in place for sufficient time for its effects both intended and unintended to make themselves apparent.

WHO CARRIES OUT A PROGRAM EVALUATION?

An evaluator is someone who is demonstrably qualified either by training or experiences (preferably both), to design and conduct a study which will be capable of addressing the concerns expressed by the principal stakeholder(s):

The competence of an evaluator is usually judged by his/her training and track-record. Has he/she carried out previous evaluations? Were they carried out efficiently? Was he/she responsive to the needs of the stakeholders? Were the results reported on time? Were the results credible and comprehensible to the stakeholder(s)? Were the stakeholder(s) satisfied that the evaluation provided them with the information they required? It is usually counter-productive to give the responsibility for an evaluation to someone unqualified by experience or training.

SHOULD THE EVALUATOR BE INTERNAL OR EXTERNAL TO THE PROGRAM DESIGN AND IMPLEMENTATION TEAM?

An evaluator may be a member of the implementation team or from outside it. However, there are many reasons why it may be appropriate to hire an evaluator who is not part of the program production team:

- an external evaluator can be hired for his/her qualifications and expertise. Professional evaluation expertise simply does not exist on many program production teams
- an external evaluator can be brought in to do a specific job at a specific time
- an external evaluation may bring a fresh perspective to the issues being addressed by the program
- an external evaluator, because he/she is not a stakeholder, may be more credible to all of the stakeholders than someone from inside the production team
- an external evaluator may be less constrained by the interpersonal and political pressures which not infrequently build up inside a

program instruction team.

WHAT IS THE ROLE OF THE EVALUATOR?

Essentially, the evaluator provides a service to the stakeholder(s). The evaluator's role is to provide the highest quality service possible to the principal stakeholder. This means that the evaluator has the responsibility for providing information to the principal stakeholder(s), information which is:

- responsive, that is, the evaluator addresses the questions and issues raised by the stakeholder(s), NOT some private interest of his/her own
- timely, that is, the evaluator must provide the results of the evaluation when they are required by the principal stakeholder. Results, however accurate and well-presented, are useless if they arrive too late for the principal stakeholder(s) to use them
- relevant to the concerns expressed, that is, the evaluator must show the logical relationship between the information collected (test scores and statistical analyses; questionnaire results; descriptions of program activities, etc.) and the questions to which the principal stakeholder wanted answers
- credible to the principal stakeholder(s), that is, the information presented must not only be relevant, but must be plausible to the principal stakeholder(s)
- comprehensible, that is, the principal stakeholder(s) must be given information in a form that he/she finds understandable. Tables of statistical analyses are inappropriate if the principal stakeholder (e.g. a group of parents, a school board administrator, a politician, a group of teachers) does not understand them.

WHAT ARE TYPICAL PROBLEMS ASSOCIATED WITH HAVING AN EVALUATION CARRIED OUT?

Programs may typically exhibit some of the following problems.

- The stakeholder(s) initially may not be clear about what questions concerning their program they want answers to. This can be overcome by the evaluator and the stakeholder(s) working together from the outset to expose the stakeholder(s) concerns, formulate appropriate questions and demonstrate how answers to these questions will be used by the stakeholder(s).
- The evaluator may collect a vast quantity of information in the hope that some of it will be of value to the stakeholder(s). This can be overcome by ensuring that the issues to be addressed by the evalua-

tion are precise and accurately represent the interests of the stakeholder(s).

- The evaluator may view his/her task as answering questions of interest to him/her personally, but of no relevance to the concerns or interests of the stakeholders. (See previous solution).
- The evaluator may have a restricted repertoire of evaluation strategies, e.g. a hypothesis-testing model in which one program must be compared with another, or a 'correlational type' model where many important variables are ignored and the relationship between one or more variables which are relatively easy to control and one or more variables which are relatively easy to measure, is demonstrated statistically. If the stakeholder(s) want(s) to test a hypothesis concerning their program or compare their program with another or assume certain variables are irrelevant and focus on the relationship between others, there is likely to be little problem. But if the stakeholders want answers to questions which cannot be answered by these evaluation strategies, the stakeholders may end up with an evaluation report which is of little or no use to them or which they cannot understand and therefore cannot make use of. On the other hand, an evaluator familiar only with qualitative strategies may provide the stakeholder with a detailed description of what happened in the ESL classes observed, or a description of what tasks the students were observed to carry out in English and include actual examples of typical student performance at different levels of proficiency. This approach may be of little use to a stakeholder who wants to see tables of mean scores obtained by students to whom a standardized test was administered.

Potential problems of mismatch between what the stakeholder wants and what the evaluator provides must be dealt with at the planning stage of the evaluation.

HOW DO WE KNOW IF A PROGRAM EVALUATION HAS BEEN SUCCESSFUL?

A program evaluation has been successful if the stakeholder(s) acknowledge that the evaluation has:

- addressed the concerns of the principal stakeholder(s)
- answered the questions posed by the stakeholder(s) using an appropriate methodology and professional rigor to gather relevant information
- supplied answers in the form of information which the stakeholders find credible
- presented the findings in a form that the stakeholders understand

- completed the evaluation and provided the findings to the stakeholder(s) in time for the stakeholder(s) to take any action needed to make any decisions that need to be made.

By contrast we know that an evaluation has been unsuccessful if the evaluation:

- did not address the concerns of the principal stakeholder(s)
- did not answer the questions of the stakeholder(s) of using a rigorous and appropriate methodology
- provided information to the stakeholder(s) in a form they could not understand or found unconvincing
- completed the evaluation too late for the results to be used to help make the decisions that had to be made.

GIVEN THE EVIDENT COMPLEXITY OF DESIGNING AND CARRYING OUT A PROGRAM EVALUATION, IS IT WORTH DOING?

A program evaluation is worth doing if the principal stakeholder(s) show that:

- it will provide him/her with information otherwise unobtainable
- that the information *can* be put to profitable use
- that the information *is likely* to be put to profitable use.

A program evaluation is not worth doing if:

- it does not address issues of concern to the stakeholders
- generates information which *cannot* be used
- generates information which *is not likely* to be used.

A simple way to test whether an evaluation should be carried out or not is for the (principal) stakeholder(s) to complete the following sentences:

What I need to know about this program is:

With this information I will be able to:

- e.g. What I (as the program designer) need to know about this program is whether the students find it challenging or not sufficiently challenging.

With this information I will be able to revise the program to make it more/less challenging.

- e.g. What I, (the teachers' association representative) need to know about this program is whether the teachers received adequate in-service preparation in order to teach the program comfortably and successfully in class.

With this information, I will be able to decide whether no further

teacher preparation is required, or whether more in-service preparation is advisable and to take steps to see that it is provided.

e.g. What I, (the Principal of a school or the Director-General of a School Board) need to know about this program is whether there are too many or too few objectives to be covered in the time allotted in our timetable for instructions.

With this information, I will be able to decide whether to increase the instructional time allotted, or to leave it unchanged.

If the (principal) stakeholder(s) can complete these two sentences, that is, he/she/they can:

- formulate a question which has two or more possible answers
- demonstrate they want and care about the answer to the question
- demonstrate how they will be able to use the answer to the question.

Then an evaluation designed to answer that question is probably worth doing.

CONCLUSION

The logic underpinning program evaluation as I have described it is one which views the role of evaluation as the provision of a service to the principal stakeholders so that the program and the management of the program can be improved. It is a cooperative activity in which an evaluation specialist places his/her expertise at the service of the principal stakeholders and provides them with credible, understandable and useful answers to their questions.

If evaluators play this service role to the best of their professional abilities, I believe that all those who have a stake in the development and outcomes of language programs will be encouraged to embrace evaluation as an essential management tool.

(In preparing this paper, I have drawn extensively upon the work of Michael Quinn Patton, in particular *Utilization Focused Evaluation*, 1987, Sage, Beverly Hills, and a document I prepared at the request of the Ministère de l'Éducation du Québec, *Program Evaluation, What's That?*, 1987.)

THE AUTHOR

Ron Mackay's interest in program evaluation is a natural extension of his work in program design and materials development. He has participated in evaluation studies for the Education Departments of several provinces and territories and for the Federal Government of Canada. He is an associate professor of Applied Linguistics at Concordia University in Montreal. His publications include "Languages for Specific Purposes: Program Design and Evaluation" (1981 Newbury

House, with Joe Palmer ed.) and, with Maryse Bosquet, a widely used ESL program SAM on Radio 321 (Longmans). He is currently engaged in research into Native Canadians and the public school system.