

# **Educating Through Educational Evaluation: An Idealistic Response to Ronald Mackay's "Program Evaluation and Quality Control"**

**Alister Cumming**

Mackay is, as usual, the consummate pragmatist. Moreover, his paper "Program Evaluation and Quality Control" is a comprehensive discussion of the more crucial aspects of second language program evaluation. How does one respond to comprehensive pragmatism?

The tack I am compelled to take is to respond to this well-founded pragmatism with a dash of idealism. In doing so I intend to reinterpret several points Mackay has already established. I also wish to argue (perhaps over-idealistically) that evaluations of second language programs can have more beneficial, illuminating and progressive effects than are usually acknowledged (or realized). My idealistic stance is that the evaluation of educational programs is, at best, an educational task — with profound potential for the improvement of pedagogical practice.

This is not a novel position: it has been advanced recently by many well-established evaluators (Marshall and Peters 1985; Weiss 1983; Cronbach 1982; and Wise 1980). The value of this perspective is that evaluators are considered to perform more than just the business of a "service" function, supplying relevant information to those who pay for it. Evaluators who educate help people to learn, to develop themselves, and to refine their abilities to do their work more effectively.

## **Why See Evaluation as Education?**

The main part of this paper will exemplify how the evaluation of language programs can educate. It reviews several evaluation studies, showing how they provide seven kinds of educational benefits. But first a few general points are necessary to justify my particular pursuit of idealism. Why should the evaluation of second language programs be viewed as an educational activity?

First, it specifies a worthy, guiding purpose for program evaluators. A language program is evaluated so that those who participate in it can learn more about it and can decide what best to do in it. As educators, the primary responsibility of evaluators is to promote learning. This role is compatible with the conventional assumptions that evaluators should be

special kinds of scientists, assistants, or judges (Wise 1980). But it suggests a responsibility beyond that of simply endorsing existing policies, testing student outcomes, or patting authorities on the back for past performance. Encouraging as such gestures may be, they have little other value unless they lead to learning by program participants and progressive developments in program implementation. Indeed, my second reason for arguing that program evaluation be seen as educational is that periodic assessment seems necessary to foster progressive developments in almost any educational program. The work of education is complex, consequential and multi-faceted. At the same time it is prone to routine, inadequacy and ineffectiveness. Continuous development, of the kind prompted by evaluation, is necessary to enhance our understandings, offer alternative insights, and supply external verifications. Evaluations which do this are fulfilling educational functions, though certainly not all evaluations do.

Third, valuees can expect to learn from the process of program evaluation. If evaluation is conceived of as educational, then participants should approach the experience as one where they will be learning actively. Ideally, program participants might engage in the process of evaluation as actively as evaluators do — rather than fearing external judgments, avoiding unusual questions, or concealing controversial issues. If evaluation is educational, then program participants are in the position to benefit most from the process. Finally, the effectiveness of an evaluation can be assessed by its ongoing and long-term impact — its educating and educational effects. Though the guidelines Mackay has outlined — based on mutual fulfillment of an agreement between evaluators and valuees — may ensure the short-term ‘success’ of an evaluation, the long-term effect of a program evaluation is also vital. Did people learn from the experience? Are they better able to perform as a result of it? Did the evaluation have a worthwhile effect? These are questions we need to ask of an evaluation, above and beyond its contractual obligations.

## **WHAT ARE SOME EDUCATIONAL BENEFITS PROVIDED BY PROGRAM EVALUATION?**

### **Validating Educational Innovations**

Evaluation is often properly called into play when educational innovations are attempted. Does the innovation work? If so, how? With what advantages and disadvantages? In addressing such issues, program evaluation attempts to provide objective answers to questions central to program implementation or change. Evaluation aims to educate by providing arguments or judgments of merit (what are the qualities of the innovation?) and worth (is the innovation worthwhile in the present

circumstances?) Evaluation of this kind is commonly commissioned to assess novel, contentious or changing language programs and policies (Cumming 1987; in press), such as bilingual or immigrant education. In practice, however, few of the novel approaches we take to language teaching receive the kind of scrutiny necessary to confirm their merits, to identify their shortcomings, or even to understand their component elements adequately.

An instance of where such inquiry has been undertaken extensively in Canada is in evaluations of French Immersion programs. Reviewing more than a hundred evaluations conducted for different school boards and agencies, Swain and Lapkin (1982) show these studies confirming that French Immersion provides an effective alternative to conventional, mother tongue education. However, since these evaluations have mainly used test results to assess the products of these programs, it is evident that further evaluation needs to assess ongoing instructional and learning processes as well as relevant social factors — before a wholesale proclamation of validity can be made. For instance, Mackay (1981) and Spolsky (1978) propose how a variety of aspects which are (respectively) internal and external to a language program may require evaluation in order to assess the program's overall effectiveness and validity.

### **Informing Program Development**

An exemplary instance of evaluation progressively informing program development is a long-term project that I (and a team of 24 others) have been working on over the past year. The project is developing computer software and school curricula which foster “intentional” learning (Bereiter and Scardamalia 1987; in press): Computer-Supported Intentional Learning Environments (CSILE). A prototype of the program (based on principles of cognitive learning) is being piloted this year in two grade six classes in Toronto. Six researchers meet weekly with two students (designated as high and low intentional learners) each, to observe and document how they are using the program. From these ongoing case studies information is gathered on cognitive, curricular and technical problems and achievements. Findings are then conveyed to system designers, computer programmers, and the students' teachers (who are also interviewed regularly to obtain their perspectives on the program's development).

This continuous, interactive, formative evaluation serves to guide enhancements to the computer functions, classroom curricula, and students' learning. While software and curriculum units are being designed and implemented, they are simultaneously being evaluated by student and teacher users as well as researcher/observers — each of whom contributes

to decisions about further refinements and developments of the overall program. At the same time, research on optimal and ineffective student learning is being conducted.

### **Illuminating the Perspectives of Learners**

A frequent educational benefit of program evaluation is to document the perspectives of students in such a way as to enrich the knowledge of teachers. Perhaps the nicest published example of this for second language instructors is a paper by Savignon (1981) written as an open letter to a teacher who had instructed Savignon in a Spanish course. She describes the course vividly from the position of student participant — assessing the classroom events which encouraged or frustrated her, the ways she was instructed to learn, and the uses she was later able to make of this language when travelling. This “microcosmic” portrayal of an informed, adult learner of a language would enlighten anyone teaching a foreign language at a university or college. Though the paper is an evaluation of a language course, it is a compelling insight into what a language learner thinks and does in such a course. The evaluation conveys a participatory perspective which is vital to language education — but which is usually obscured in our more popular concerns for teaching methodology, research design and curriculum planning.

### **Clarifying an Educational Rationale**

Evaluation can also help educators to understand better what they already do well. Refinements can be made to educational practice through clarification of purpose. Such was the case in an evaluation I worked on (Cumming and Burnaby 1986) studying a cooperative Chinese/Canadian program bringing Chinese professionals to Canada for further education at different businesses, agencies and universities. The program’s mandate had been to accommodate diverse learning needs for a great variety of professionals in unique circumstances. This called for a flexible educational design, one that the program organizers had arrived at implicitly by drawing on different elements from conventional models of adult education to suit individual purposes for study. In the process of evaluating the overall program, however, it became apparent that the philosophies of some of the educational models contradicted one another, producing inevitable differences in the expectations of learners, instructors, or program organizers. For instance in several cases, the roles of (1) “trainer-trainee” assumed for short-term training or upgrading courses tended to conflict with aims of (2) cooperative, cross-cultural exchange or long-term institution building, which require relations of equal status among partici-

pants. The evaluation study was able to identify instances where the existing program models worked most effectively for cooperating participants. On the basis of this analysis a more distinct, pedagogical philosophy (based on reflective inquiry and observation directed at long-term applications of knowledge) was recommended to guide the design of future programs.

### **Proposing Ethical Criteria**

Evaluation can also educate educators by proposing criteria to guide the ethics of their work. Noteworthy instances of this appear in evaluations by Hayhoe (1986) and McLean (1986), also related to recent Canadian/Chinese educational exchanges. Hayhoe's studies of a large number of programs implemented by Western institutions and agencies in China have prompted her to develop a set of principles for assessing the equity and effectiveness of these programs. Her principles center on the notion of social mutuality — whether there are just, mutual benefits for cooperating participants. In particular, are equity, autonomy, solidarity and participation achieved cross-culturally during the implementation of such programs? Evaluation can establish the extent to which this may be true for different participants, in the short and long terms, and for different aspects of activity.

### **Bringing to Light Social Inequalities**

Cummins (1984, pp. 19-65) provides an example of how the evaluation of decisions taken in an educational program can help us to understand, and one hopes, ameliorate certain unfair biases in the practice of schooling. Cummins reviews a large number of assessments made by teachers and psychological consultants who had tested and made referrals for ESL students in an Alberta school board. Quantitative and qualitative analyses show the abilities of students from minority language backgrounds had often been misinterpreted on the basis of their test results and classroom behavior. Decisions were made which were insensitive to: the cultural and linguistic biases underlying IQ (and other) tests; the conditions which would promote cognitive and academic development for minority language students; and the values of families wishing to maintain the use of their native languages in their homes. In evaluating these circumstances certain inequalities underlying common practice are exposed; ways of better educating practitioners in these matters become evident. Evaluation is a process for educating ourselves and our colleagues about what we commonly do.

## **Appreciating the Art of Educating**

My final example is of evaluation helping us to understand better the “art” of education — the virtuoso performances of people in time and place characterizing the finest achievements of education. Eliot Eisner (1979) has urged evaluators to replace the quasi-scientific measurement which has dominated educational research with an approach to evaluation which follows principles of connoisseurship, as in art or literary criticism. Knowledgeable evaluators prepare appreciations of educational experiences, showing how qualities of performance, context and beliefs interact to create significant (or insignificant) events. Such evaluations are written, like inspiring criticism, to evoke the experience of participating in the object of study. The valuable elements of an educational program are glorified (if appropriate) so that educators, participants and the public can appreciate them fully.

This approach to evaluation has not yet been developed in second language education, though Mueller (1983) makes gestures in its direction in a study of foreign language teaching in a U.S. school. But the approach promises to offer a means for evaluation to demonstrate (in a way that would only be obscured by reductive experimentation) how and where language teaching excels, enriches and ennobles.

## **CLOSING REMARKS**

What I have called my idealism is really an expression of hope about what evaluation might do for second language education. But idealism only contemplates the relation of theories to the world. It is pragmatism — intervening to do things — in the way Mackay has defined evaluation, which must inform the practical work of evaluation and determine its benefits. To narrow gaps there may be between the idealistic and pragmatic perspectives, let me close the discussion by pleading my beliefs more directly and pragmatically.

I want to encourage people involved in language education to use a broader, productive conception of program evaluation — with the aim of educating ourselves more thoroughly in the work we usually do — in order to improve our professional abilities. Let us not simply accept that a certain teaching approach “works”, a certain classroom routine is conventional, a certain textbook has ministry approval, a certain learning task is sufficient, a certain curriculum is prescribed, a certain test demonstrates one kind of validity, a certain policy is mandated, a certain community has specific learning needs, a certain research finding suggests something, certain students are motivated or unmotivated learners, or even that a certain evaluation project has predetermined outcomes.

Let us study, reflect on, and assess these things. Evaluate them — conscientiously, formally and informally, and productively — as part of our ongoing responsibility to learn how to do language education more effectively. Whether it be part of an organized evaluation study or daily educational practice, I would hope that everyone, in their particular pedagogical settings, can work toward practices which strive (as I have suggested above) to: validate educational innovations, inform program development, illuminate the perspectives of learners, clarify educational rationales, adopt ethical criteria, bring to light social inequalities, and appreciate the art of educating. These are educational lessons I would hope exemplary program evaluations are able to teach us.

---

## REFERENCES

- Bereiter, C. and Scardamalia, M. (1987). An attainable version of high literacy: approaches to teaching higher-order skills in reading and writing. *Curriculum Inquiry* 17(1), 9-30.
- Bereiter, C. and Scardamalia, M. (In press). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.) *Cognition and instruction: issues and agendas*. Hillsdale, N.J.: Lawrence Erlbaum.
- Cronbach, L. J. (1982). *Designing evaluation of educational and social problems*. San Francisco: Jossey-Bass.
- Cumming, A. H. (1987). Evaluations and developments of foreign language education in China. *Canadian and International Education* 16(1), 211-220.
- Cumming, A. H. (In press). What is a second language program evaluation? To appear in *Canadian Modern Language Review*, 1987.
- Cumming, A. H. and Burnaby, B. (1986). Final Evaluation Report on the China/Canada Human Development Training Program: Models of Organization and Pedagogy and Their Potential Impact. Unpublished report submitted to the Canadian International Development Agency. Toronto: Ontario Institute for Studies in Education.
- Cummins, J. (1984). *Bilingualism and special education: issues in assessment and pedagogy*. Clevedon, Avon: Multilingual Matters.
- Eisner, E. W. (1979). The use of qualitative forms of evaluation for improving educational practice. *Educational Evaluation and Policy Analysis* 1(6), 11-19.
- Hayhoe, R. (1986). Penetration or mutuality? China's educational cooperation with Europe, Japan and North America. *Comparative Education Review* 30(4).
- Mackay, R. (1981). Accountability in ESP programs. *ESP Journal* 1(2), 107-121.
- Marshall, J. and Peters, M. (1985). Evaluation and education: The ideal learning community. *Policy Sciences* 18(3), 263-288.
- McLean, L. (1986). Overview of the Development of the China/Canada Enterprise Management Training Centre at Chengdu, Sichuan. Unpublished report submitted to the Canadian International Development Agency. Toronto: Ontario Institute for Studies in Education.
- Mueller, M. (1983). The tower of Babel in Libertyville. *Daedalus* 112(3), 229-247.

- Savignon, S. (1981). A letter to my Spanish teacher. *Canadian Modern Language Review* 37(4), 746-750.
- Spolsky, B. (1978). A model for the evaluation of bilingual education. *International Review of Education* 24(3), 347-360.
- Swain, M. and Lapkin, S. (1982). *Evaluating bilingual education: A Canadian case study*. Clevedon Avon: Multilingual Matters.
- Weiss, J. (1983). Curriculum commonplaces and evaluation counterparts. Paper presented at the Annual Convention of the American Educational Research Association, Montreal, April 12.
- Wise, R. I. (1980). The evaluator as educator. *New Directions for Program Evaluation* 5, 11-18.

### THE AUTHOR

Alister Cumming works in the Faculty of Education at McGill University. He is interested in process of thinking, learning and writing in Second Language curricula and instruction. He has worked at a variety of universities in Canada, most recently the Ontario Institute for Studies in Education.