I.Œj 0'Cnf gtuqp.'C0'Dgtgwc'*gf u0<*Evaluating Second Language Education.*

Harold Ormsby L.
UACPyP/CCH-UNAM

What Alderson and Beretta have done with this book is to bring together and to print information that has been dispersed in the "oral tradition" of second language education evaluation for some time now. If that were all it did, it would be enough to justify the authors' and editors' efforts. However, they have gone beyond a simple recounting by focussing on theoretical and methodological questions and by suggesting a set of guidelines that can be expected to become a significant contribution to the field. They have also made it uncomfortably clear that virtually all second language program evaluations have been failures precisely because they have neither paid sufficient attention to theory and methodology nor learned from others' experiences.

The book is divided into three parts. Beretta begins the first part with an overview of L2 program evaluation over the last twenty-five years. He then moves on to a brief discussion of the broad trends and what has been learned in non-language-centered educational evaluation since the 1960s. This part ends with Beretta's predictions about what the field of language education evaluation will be like in the 1990s. In this regard, he feels that the field will (1) no longer pose unanswerable research questions, (2) stop aspiring to advance the separate field of language learning theory, (3) be taken more seriously and, hopefully, manage to get program evaluation included in syllabus designs from the very beginning, (4) learn how to encourage and carry out negotiations among clients, evaluators and stakeholders, how to distinguish between the feasible and the impossible and how to prioritize, (5) master a number of quantitative and qualitative research methodologies because "no single design can hope to address the great diversity of questions that are asked about programs," (p. 20) and (6) learn how to translate academic findings into the language of policy. He hopes that evaluations themselves

"will shed light on the nature of program design, development and implementation" and that "reflexive evaluations, such as those reported in this book, will offer insights into the nature of evaluation" itself (p. 20).

The second part of the book consists of eight "Case Studies of Current Practice" and brief postscripts by the editors. In the first study, Alderson and Mike Scott report experiences in the English for Specific Purposes in Brazilian Federal Universities project. They bring out the rather absurd nature of the JIJOE ("Jet-In Jet-Out Expert") and the advantages of "participatory evaluation," which relies heavily on the participation of insiders. In the second study, Bryan Lynch discusses the Reading English for Science and Technology (REST) project, a cooperative venture between the University of California, Los Angeles, and the Universidad de Guadalajara. He does not report any conclusions about the project itself; rather, "the focus and conclusions of this study [are] on the use of quantitative and qualitative data and their associated methods of analysis in the evaluation of language teaching programs in general" (p. 61). The third study deals with the question of supposedly independent, unbiased evaluators who are called in (at the last moment) by policy-makers. Rosamond Mitchell briefly describes a program to maintain Gaelic in the Western Isles of North-West Scotland and shows how language-program evaluation can get embroiled in the stormy seas of local and national politics, dirty tricks and all. In the fourth study, Adrian Palmer reports the discussion and negotiation processes that led to the evaluation design that came to be part of an experimental German course that was set up (in 1985-86) to put aspects of Krashen's theory of second language acquisition to the test. Here, one is given a glimpse of how a particular team of evaluators went about facing dilemmas and finding solutions.

In the fifth study, "Program-defining evaluation in a decade of eclecticism," Steven Ross brings out the fact that under a *laissez-faire* approach to program administration, each teacher comes up with his or her own syllabus and uses his or her own favorite methods and procedures. "The classroom observation data gathered in this study demonstrate how a longitudinal approach to evaluation of intact classes can lead to a limited set of testable hypotheses about product differences" (p. 192). In the sixth study, Assia Slimani faces the fact that "learning outcomes are not necessarily the reflection of the teacher's plan since, in the process of accomplishing instructional objectives, interactive work takes place... and leads to a whole range of [perhaps unexpected] learning opportunities" (p. 197). In this case, the focus is on research techniques and not on program evaluation *per se.* However, "anyone thinking of adopting this type of inquiry in an evaluation," according to the editors, "can turn to Slimani's account and consider how they might overcome some of the problems she highlights" (p. 221). In the seventh case study, "Moving the goalposts: project evaluation in practice," Hywel Coleman gives a frank, detailed report of his experiences when he was posted by the UK Overseas Development Administration and the British Council to a Key English Language Teaching (KELT) project. His points are that projects change as various aspects of reality impinge on them and that, therefore, it is very difficult to decide whether they have been successful or not.

In the last case study, Beretta takes up the well-known Bangalore Project with a "warts-and-all" approach. The editors take the opportunity to make their attitude known:

> [M]ost evaluation reports give the impression of a flawless design, problemless administration and the production of a non-controversial report. Yet most evaluations are simply not like that; in the real, messy, political, ever-changing world of education, educational politics and personalities, compromises are the order of the day, where ideal designs for evaluation studies have to be sacrificed at the altar of undefined aims, where hopes and fears abound, and where evaluation methodologies and foci are arrived at all too often ... in a rush, without deliberation or negotiation and in some confusion. Hindsight is not only the most exact science; it is also the discipline best able to give the impression of order, rationality and careful planning (p. 272).... [Evaluation is important, and needs to be taken into account from the beginning of a project; yet at the same time it is not an exact science; it is exposed tp accusations of interference, lack of understanding of circumstances, and partiality of method and interpretation ... (p. 273).

With Part III, Alderson, while denying that there ever will be One Best Way of conducting evaluations and that they will ever reveal The Truth, attempts to "provide would-be evaluators with guidance on how to go about designing and implementing an evaluation" (p. 274). He deals with the WH questions (why, who for, who does it, what, when, how, how long) of evaluation, then goes on to such topics as "Initial negotiations," "Deadlines, deliverables and dust: What happens to an evaluation report?," "Implementing an evaluation," "Interpretation and reporting" and, finally, "Using and evaluating evaluations." This section is definitely not a recipe book; rather, it is a pantry of items that might go into a very complex thinking process. Of equal importance, there is a great deal of well-founded advice.

If it could be done, I would like to see this book made required reading for every professional who is even flirting with the idea of developing a taste for evaluation, and more so for any who have already developed it. Unlike people who love Law and sausages, those who love evaluation *must* know how it is made.