Evaluating Second Language Education
edited by J. Charles Alderson and Alan Beretta
Cambridge: Cambridge University Press, 1992.

Reviewed by Desmond Allison
The University of Hong Kong

Evaluating Second Language Education is written primarily for language teaching professionals who are likely to work as (or with) evaluators of language teaching programmes. The book presents a short editorial introduction, opening and closing chapters by one of the editors, and eight 'state-of-the-art' evaluation studies, each with an editorial commentary. In the opening chapter, 'Evaluation of Language Education: An Overview', Alan Beretta contrasts the lack of attention given to evaluation studies in language programmes with the wealth of work in other areas of education. Charles Alderson's closing chapter, 'Guidelines for the Evaluation of Language Education', attempts without prescriptivism to systematise practical issues that will arise during an evaluation. The book as a whole offers reflective experience as a way to "contribute to an understanding of the discipline of evaluating second language education" (Editors' introduction, p.2).

The eight case studies are: Insiders, outsiders and participatory evaluation (Alderson and Scott); Evaluating a program inside and out (Lynch); The 'independent' evaluation of bilingual primary education: a narrative account (Mitchell); Issues in evaluating input-based language teaching programs (Palmer); Program-defining evaluation in a decade of eclecticism (Ross); Evaluation of classroom interaction (Slimani); Moving the goalposts: project evaluation in practice (Coleman); What can be learned from the Bangalore Evaluation (Beretta). The projects are set (respectively) in Brazil, Mexico, Scotland, the U.S.A.,Japan, Algeria, Indonesia and India. English is the second language in most cases, but Mitchell's study concerns Gaelic-English bilingualism, and Palmer's study examines the learning of German in an American university. Five studies are set in universities and range from first-year undergraduate teaching to teacher education; the other studies are those by Mitchell (primary education), Ross (junior college) and Beretta (secondary education).

My response to the case studies was to skim-read the introductions and closing parts, returning to the chapters selectively according to interest, and to focus on issues and "lessons"that might be applied in my own working circumstances. Some of these were the relationship between programme evaluation and teacher development (Alderson and Scott); quantitative and qualitative evaluation (e.g. Lynch); problematic relations between evaluations and policy development (Mitchell; Coleman); ways to conduct rigorous and interesting classroom-centred research (Slimani). I also found the editors' contributions helpful in suggesting connections between the contextualised case studies and their possible applications to other circumstances. I shall keep going back to the book.

An earlier review of Alderson and Beretta by Davies (1992) raises some important questions about current approaches to language programme evaluation. Indicating courteous scepticism over insider involvement in evaluations and over a place for qualitative analysis, Davies asks whether "an evaluation necessarily means telling the reader everything" (1992, p.208). The danger of overkill in qualitative studies is real, and a fully comprehensive account of a language programme must remain unattainable. Yet one can avoid the excesses of enthusiasm without total abstinence. To avoid becoming submerged in unusable qualitative data, Lynch (p.93) specifies the need for an explicit purpose when drawing upon and categorising information. As sources of such data, Lynch mentions journal entries, daily logs, observations, meeting notes, questionnaires and interviews (p.69), while his coding categories (p.77) include goals, processes, events, settings, participants and outcomes. Despite Davies's misgivings

on the point, qualitative data can offer important insights for an evaluation. For example, the frequent mismatch between teaching plans and learning outcomes (Slimani) might simply go unnoticed if evaluations were carried out solely by outsiders concerned with performance measures.

Davies's main problems with the book, however, concern its editorial stance and its case study approach. The editorial commentaries on the case studies involve some repetition of themes, and Davies finds a need for a critical review of second language evaluation that is separated from accounts of practice. Above all, Davies finds the major theme "dispiriting" in its rejection of the possibility of objectivity in programme evaluation.

The adoption of a case study approach is presented by the editors themselves as a principled choice. Alderson and Beretta in their introduction (pp.1-3) describe second language programme evaluation as requiring "ad hoc investigation that appears to have no tradition that can be appealed to", in contrast with evaluation studies in other social and educational spheres (a tradition drawn upon in Beretta's overview chapter). The use of case studies is intended to capture experiences and problems of evaluators in particular settings, and to avoid unreal "cleaned-up" accounts of the evaluation process. While commenting that all eight case studies are "firmly grounded empirical enquiries", Alderson and Beretta observe that "at this stage of our development, the history of an evaluation is probably more important than its findings". They suggest the value of learning from accounts of decision-making processes during evaluations.

As a mode of knowing, the case study approach appears better respected and more highly valued in applied (educational) linguistics today than was often the case before the 1980's. Case studies include an important element of narrative (explicit in Mitchell's title). Hymes and Cazden (1980) argue for greater recognition of narrative thinking, presenting it as a legitimate but greatly undervalued way of exploring and conveying knowledge.

If anything, a case study approach tends to privilege other forms of knowing than those most closely associated with technical rationality and efficiency. Work in the ethnographic tradition has encouraged case studies, allied to research practice whereby the observer seeks to formulate descriptive and evaluative categories used by participants themselves, rather than imposing a preconceived framework of analysis; van Lier (1988) among others has brought this tradition firmly to bear on the second language classroom. Careful exploration of experience in particular contexts also appears crucial in the development and validation of "local forms of knowledge" called for by Pennycook (1989, p.613).

While the case for in-depth explorations of particular situations seems compelling in this light, there remains a need for attempted generalisation if accounts of various evaluation studies are to "contribute to an understanding of the discipline of evaluating second language education" (editors' introduction). This brings us to questions of a disciplinary framework, and of the status of knowledge claims.

Candlin observes that much second language research lacks an overall theory of discourse and a social orientation, and consequently produces "underdetermined analyses which may act to trivialise the complexity of human social interaction" (Candlin 1990, p.474). Although his concerns are mainly with the limited vision of narrowly-focussed studies in the positivist tradition, Candlin also points out the danger of producing only "some mere listing of unmotivated and unwarranted subjective meanings" (p.479) in seeking to acknowledge the place of subjectivity and of differing viewpoints when undertaking data collection and analysis.

Davies's doubts about qualitative analysis, and perhaps about a case study approach to the development of a discipline, here find an unexpected echo. However, the responses of Candlin and Davies towards subjectivity and differing viewpoints could scarcely be more divergent. Favouring ethnographic rather than experimental accounts, Candlin argues that experimental work has a place within a more comprehensive and self-questioning or "reflexive" approach towards the use of research

procedures in applied linguistics. Such a reflexive approach to data gathering and treatment appears highly relevant to language programme evaluation (Beretta, p.20; Alderson, p.299). It appears able to handle multiple viewpoints, and to give weight both to subjective judgements and to an aspiration towards generalisable knowledge.

In seeking to encourage a disciplinary perspective, Alderson and Beretta do not themselves propose any overall theory or framework for the evaluation of second language education. The nearest they come to a model is in Alderson's "guidelines" to the evaluation of second-language education programmes, which aim

> ... to offer suggestions about how to set up and carry out evaluations in any given setting. The section attempts to provide a balanced overview and discussion of the issues encountered in the evaluation of second language education ... (and) to furnish insights into the nature of evaluation in a way that is intended to provide practical guidance to would-be evaluators (Alderson and Beretta, pp.2-3).

An outline model of the stages of an evaluation is then used for expository convenience. Alderson's account follows "the usual and logical stages in the conduct of an evaluation: planning, implementing, interpreting, reporting, using, evaluating" (p.274). (For more detail, see my article in this issue.)

Using upper-case letters to stigmatise claims that he considers excessive, Alderson rejects the notions that there might be "One Best Way" to conduct an evaluation and "One Truth" to reveal. Purposes, projects, people, timescales and resources will all affect the choices to be made during an evaluation. Different interpretations will need to be constructed and presented from (and for) different viewpoints. Alderson reassures us that this "emphatically does not mean that 'anything goes': it is essential that evaluations be conducted in a principled, systematic and explicit manner", but he also insists that "No evaluation is ever objective...The best we can hope for is pooled intersubjectivity and reduced or neutralised partiality" (Alderson, pp.274-275).

Reacting to such observations, Davies (1992) bemoans what he describes as "a strong editorial undertow... offering only a dispiriting doubt as to whether evaluation is ever possible". Dismissing preoccupations over the partial and often partisan nature of judgements, and the impossibility of an objective stance, as "an extreme observer's paradox position", Davies adds "one longs in the editorial commentary for the sort of sprightly kick Dr. Johnson offered to the anomic gloom of Bishop Berkeley" (Davies 1992, p.207).

Potted philosophy is dangerous - a comment that could promptly backfire on this writer. However, Davies's allusions to Berkeley notwithstanding, a conviction that an "objective" stance is unattainable surely need not lead us to reject external reality, or to renounce all prospects of socially constituted knowledge. For example, Lakoff (1987) offers a resilient and cheerful optimism about human capacity to attain knowledge, while at the same time arguing against "objectivism" and in favour of "experientialism" as the basis upon which our knowledge of the world has necessarily to be composed. To suggest that "reality" is observed, and constructed, from some point of view is (in this view!) to accept the conditions by which reality can be known, and not to despair of its existence.

The problem for objectivism, which at least has serious implications also for a positivist belief in "objectivity", is that reality cannot be known and conveyed independently of some point of view. The ideal "God's eye" view may be the ultimate synthesis of all views, rather than the one perfect vantage point. Such a synthesis is humanly unattainable - which also offers a principled reason to keep enthusiasm for qualitative analysis within manageable bounds.

What might an experientialist view of reality imply for the evaluation of educational programmes?

136

External and educational realities exist, but different elements of reality will be more or less salient, and more or less closely related, according to the perspectives from which they are considered. Any reasonable degree of consensus among people will reflect a synthesis of views that is itself put forward within a shared and socially constructed perspective. A successful evaluation report, then, will offer an account that gives sufficient heed to the perspectives of different "stakeholders" in a project to be acceptable to these groups or their representatives, with acceptability and success being matters of degree, and related to users in contexts. Above all, a successful report will actually be used.

To assert that second language programme evaluation is a practical activity is not, however, to consign it wholly to the realm of the arbitrary, the makeshift and the convenient. An evolving "discipline" of second language evaluation seeks to establish standards of practice and to afford its own socially constructed perspective, wider in space and time yet narrower in "professional" focus, upon this area of activity. Alderson and Beretta's book brings us closer to such a discipline. In so doing, *pace* Davies, *Evaluating Second Language Education* also offers hope.

References

Candlin, C.N. (1990). What happens when applied linguistics goes critical? In Halliday, M.A.K., Gibbons, J. and Nicholas, H. (Editors), *Learning, Keeping and Using Language (Vol. II): Selected Papers from the 8th World Congress of Applied Linguistics, Sydney, 16-21 August 1987.* Amsterdam: John Benjamins.

Davies, A. (1992). Review of Alderson and Beretta (Editors). *Language Testing, 9(2),* 207-209.

Hymes, D. and Cazden, C. (1980). Narrative thinking and story-telling rights: a folklorist's clue to a critique of education. In Hymes, D. (1980), *Language in Education: Ethnolinguistic Essays,* 126-138. Washington, D.C.: Center for Applied Linguistics.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What categories reveal about the mind.* Chicago: University of Chicago Press.

van Lier, L. (1988). *The Classroom and the Language Learner: Ethnography and second-language classroom research.* London: Longman.

Pennycook, A. (1989). The concept of method, interested knowledge and the politics of language Teaching. *TESOL Quarterly, 23(4),* 589-618.