Designing evaluation into educational change processes

Christopher Tribble

In this paper I shall discuss the integration of insider-managed baseline evaluation into project and programme design in educational change processes. Drawing on examples taken from recent experience of ELT projects in Central Europe and China, I shall outline some of the benefits of using insider-managed baseline studies—and some of the problems which can arise when this kind of study has not been carried out at the beginning of a change process. Throughout the paper I shall stress the importance of including implementation teams in evaluation processes, and will show how baseline studies can be used as a means for ensuring that this inclusion takes place.

Evaluation, baseline studies, and stakeholder analysis

Trends in evaluation

A well-established professional consensus on approaches to evaluation in development and education is now being more systematically applied to the design of English Language Teaching projects and programmes (Alderson and Scott 1992, Weir and Roberts 1994, Murphy *et al.* 1995, Weir 1995). This consensus on the kinds of evaluation which may be required in relation to a change process, and the reasons for, and benefits from such evaluations, are summarized in Table 1.

Table 1

Category	Timing	Main purpose	Main benefit
Baseline	• prior to project approval • at project inception • at mid-stage	• assures future evaluation • contributes to ongoing and forward project planning	• makes it possible to assess the need for the project • establishes a basis for future evaluation • provides important milestones in projects with long duration
Developmental/ Formative	• during project implementation	• improves implementation	• contributes to effective implementation • improves project team performance • permits reasoned changes to project design in response to changing circumstances • enables effective communication with interested stakeholders
Summative/ Impact	• on project completion • some time after completion	• accountability • learning	• provides information for funding partners and other interested authorities • allows lessons to be learnt for application in future projects

Not only is there this consensus amongst professionals on the kinds and purposes of evaluation; evaluation itself is being given much more of a central role in the work of many development organizations.[1] In such an environment, it is not surprising that evaluation is also taking a more central position in the work of implementing organizations, such as the British Council.[2]

Alongside this increasing demand for an evaluation process which offers 'greater accountability, better decisions, and better sharing of lessons learned' (OESP 1997: Chapter 2), there is also an understanding that it is not only *what* you evaluate which matters. The *who* and the *how* of evaluation are also of critical importance if it is to provide information which will be useful to sponsors, implementing agencies, and beneficiaries (King *et al.* 1987, Weir 1995). This has resulted in a recognition of the importance of ensuring greater involvement of those who are 'doing' the project in the process of evaluation. By combining insider management and involvement with outsider consultation and moderation of evaluation, a double benefit accrues:

- the need for accountability is met, and
- it is possible to get better projects implemented by more professionally effective teams.

In this paper I shall also be arguing for the use of insiders in evaluation processes. I shall situate this discussion by focusing on the value of such involvement in a crucial, and still-neglected aspect of evaluation: the baseline study. In the next sections I shall outline some of the uses of baseline studies, before going on to give examples of situations in which insider-managed baseline studies have, and have not, been used.

Baseline studies

A baseline survey is a snapshot of the context in which an educational change process is going to take place. Importantly, such a survey need not restrict itself to classroom activities or examination results: it can also take into account policy questions, the roles of other suppliers, and the long-term interests of employers. Thus, a baseline survey might cover such areas as:

- Statistical information (published sources)
- National experience (published sources/activities by other agencies in the same area)
- International experience (published sources)
- Appraisal of key institutions (resources/staff/practices/policy)
- Interviews with key stakeholders[3] (e.g. Ministry/LEA/college/university/training organizations/specialists)
- Surveys of parents—needs and attitudes
- Surveys of students—needs and attitudes
- Surveys of employers—needs and attitudes

- Assessment of funding available (national)
- Assessment of funding available (other donors)

Alderson has made the following comments on the value of baseline evaluation:

> If the evaluation is intended to estimate the impact of a project, then it is desirable to attempt to establish what things were like before the project began. This necessitates the gathering of *baseline* data. Such data will need to relate to the predicted outcomes of the project . . . ' (Alderson 1992: 287).

However, baseline data collection does not have to be restricted to the beginning of the project. Baseline studies can also be undertaken at:

> 'the pre-planning stage, [. . .] at the planning stage, [. . .] and at some point in a project's implementation' (Murphy *et al.* 1995: 7).

I will refer to these different forms of baseline research as:

- **pre-initiation** baseline studies (a baseline report prepared before a project has been approved, local agreements finalized, and budgets allocated)
- **post-initiation** baseline studies (a baseline report prepared in the opening phase of an approved project)
- **milestone** baseline studies (a baseline report prepared at the end of a significant stage in a long project)

Each category of baseline study has its own value, and each can play a significant part in project design and implementation. In the rest of this paper I shall discuss the value of baseline evaluation and the ways in which insiders can and should be involved in its design, the development of data collection instruments, the collection and analysis of data, and the final dissemination of these results in reports which will be used in policy formation, formative review, and summative evaluation.

Examples and case-studies

Post-initiation baseline evaluation

In this paper there is only space to discuss one kind of baseline study, so I have chosen to focus on post-initiation baseline research, as it is in many senses the default form. Post-initiation baseline research takes place *after* a project has been approved and a budget has been allocated. After the appointment of a project manager, it can be seen as the first input to the project. Apart from its function in later evaluation, the whole idea of the baseline research is that it should offer the implementation team an opportunity to check the original project conception against closely-observed reality. The extent and focus of a post-initiation baseline study will depend on the scale and scope of the project (Bray and Luxon 1998).

Post-initiation baselines have been used in several projects in Eastern and Central Europe managed by the British Council, where the Institute for English Language Education at Lancaster University has provided

training and consultancy services. In the following section I shall outline the benefits obtained from using baseline evaluation in a project which ran in the Baltic countries of Estonia, Latvia, and Lithuania between 1993 and 1997–8.

The Year 12 Projects

Aim: The 'Year 12 Project' was planned as a set of three-year examinations reform projects which had the aim of providing the Baltic countries of Estonia, Latvia, and Lithuania with a means for developing school leaving examination systems capable of harmonizing with European Union institutions. The main outputs for the projects were to be:

(a) three new centrally-marked English Language examinations for school leavers

(b) appropriate systems and institutions for developing, producing, distributing and administering, and marking the examinations.

The intention was that these new examinations and institutions would provide a model which could be applied in the development of examinations in other subject areas.

Outcomes: The project was completed successfully in the academic year 1997–8, and achieved its aim. Each country has its new Year 12 examinations, and appropriate institutions have been developed, in the form of national examination centres, often employing former members of the Year 12 teams, which can carry forward the work of examination reform in other core curriculum areas. What made the Year 12 project special was the way in which three British institutions—the British Council, Lancaster University IELE, and the Scottish Examination Board (SEB, now the Scottish Qualifications Authority, SQA)—were able to work with national governments, the teaching professions, and the project teams (which the British Council recruited in partnership with the three Ministries of Education), to elaborate three different but harmonious projects, each of which was insider-owned and managed. The British Council English Language Teaching Consultants in the three countries provided co-ordination and support, IELE provided training, experience, quality control, and encouragement, and the SEB made available insights into the professional issues associated with managing a centrally-marked national examination. However, although they had all this support, it was the project teams (most of whom had been 'ordinary' school teachers before the project began) who developed the examinations and their associated institutions.

The baseline studies: The baseline reports which each team wrote at the commencement of the project were central to the overall project design.[4] They were the result of a major research exercise undertaken at the beginning of the project, which complemented small pre-initiation baseline studies prepared by the British Council ELT Consultants in each country. Although the Baltics' baseline studies were not perfect, they provided a number of significant benefits:

- Each team received an introduction to the kinds of research methods they would need later in the project. This enormously enhanced their self confidence—and also their future capacity to undertake research-informed examination development.
- Each team came to a full and realistic understanding of the huge task which they faced.
- The baseline reports provided the only authoritative accounts of the pre-project environment—and underscored the urgent need for reform of the examination system for modern language teaching. As we shall see below, in spite of their limitations they constituted a sound basis for subsequent evaluation.
- The baseline report provided the starting point for the individual implementation plans of each project.
- Preparing the baseline study gave the project teams the grounding, and, crucially, the confidence they needed in order to take responsibility for project development. By the time they had completed training, research, and writing up, they had become the national experts in the problem they were trying to solve!

Baseline studies in evaluation

So far we have seen that building baseline evaluation into the design of the Year 12 projects, and ensuring there was maximum insider involvement in this process, had many benefits over and above the contribution the studies would make to later evaluation. But what of the evaluation of the projects? Given the kinds of aims the projects had (to bring about new examination systems for English language education), is it not the case that the only evaluation required was to see if the projects had achieved their desired outcomes? As new and appropriate Year 12 examinations and examination systems for English are now in place in Estonia, Latvia, and Lithuania, why bother to look back at how things were in the past?

There are several answers to these questions. The first is that evaluation is not only concerned with accountability (or, worse still, with a punitive apportioning of blame for failure!)—it is also concerned with understanding success. So the baseline data would have been required whether the project had succeeded or failed. A second answer can be found in the very success of the projects in difficult circumstances. Although the examinations and exam systems are not perfect, by and large they are acceptable, and to a great extent this is because evaluation was built into the project design—and the starting point for this was the development of the baseline reports.

Insider-managed evaluation was integral to the success of the Year 12 projects, and was complemented by a second innovation—the development of systems of Process and Content Quality Assurance (QA). Process QA was provided by the three British Council ELT Consultants based in the Council's national offices.[5] The ELT Consultants ensured that reports were delivered on time and to standard, that resources were

available, and that communications were assured. Content QA was provided by an expert 'outsider' consultant from the Department of Linguistics and Modern English Language in Lancaster University, who was recruited in order to ensure a continuing external professional view on the project. At the beginning of his involvement the initial baseline studies were the main point of reference. As the projects developed, outcomes were subject to continuing constructive scrutiny, and the project teams grew in their ability to reflect on their own practices. Thus the original baseline studies provided the starting point for the projects, while ongoing professional evaluation provided by an informed outsider, working *with* the insiders responsible for implementation (who were also directly involved in the development of research questions and data collection) ensured that plans were adjusted in relation to need, so that the final aim was achieved in spite of changing circumstances.

All this has meant that throughout the project it has been possible not only to trace the elaboration of the examinations, but also to identify other less spectacular changes in the project environment, including changes in classroom practice and changes in teacher attitude towards the examinations. The baseline reports contained the picture of the *status quo ante*. Ongoing Process and Content QA ensured the collection of formative evaluation data, and provided the teams with an indication of how they were doing in relation to the initial baseline and successive external Content QA reports. The end result was that by the time the project was completed, its overall success had, to a large extent, been assured by the regular review of policy documents and implementation plans in the light of systematic insider-informed evaluation. In addition to which, of course, the existence of the baseline studies means that future impact evaluation—perhaps three or five years down the road—can be meaningful.

Managing a project in the absence of adequate baseline data

The next project I shall discuss began in China in the latter part of the 1990s, but ran into serious problems by the time of its mid-term evaluation. The project director and the implementation team had become aware that there were problems, but, in the absence of baseline data there was insufficient information to help them identify exactly what these problems were. My evaluation of the projects indicated that it was premised on several wrong assumptions about the project environment—assumptions which insiders would not have made—and that the lack of baseline data had made it difficult, if not impossible, for the project team to identify the kinds of obstacles that they would face as they attempted to implement a planned second stage.

The project: The project, which had been designed by an external consultant, was built around a cascade to be initiated by the training of a cohort of teacher-trainers in the major Normal University in the province. The Logical Framework[6] left column outlined the Goal/Purpose/Outputs/Inputs (Activities)(Table 2).

Table 2

Goal: To contribute to the teaching of English in middle schools across the Province by up-grading the training of teaching personnel
Purpose: 1. to train designated key resource persons as trainers who will in turn train in-service teachers and other teacher trainers 2. to develop a 4-year undergraduate programme so as to provide a supply of competent middle-school teachers for the Province
Outputs: 1. a project team able to set up and deliver trainer training, and to set up teacher training courses 2. an initial body of qualified teacher trainers 3. a programme for teacher training courses based on pilot courses 4. a 4-year undergraduate programme at the Provincial Teachers' University
Activities: 1. trainer training courses 2. teacher training pilot courses 3. ongoing staff development at the Provincial Teachers' University 4. curriculum development at the Provincial Teachers' University

Problems: The problems clustered around the first of the Project Purposes:

> 'to train designated key resource persons as trainers who will in turn train in-service teachers and other teacher trainers'

. . . and the second project output:

> 'Output 2: an initial body of qualified teacher trainers'

Unfortunately, these planned project components were premised on the false assumption that ordinary school teachers trained through the project would be in a position to become INSET tutors for schools in their region. While there is nothing wrong with notions of mentoring and school-based teacher development—they have been successfully implemented in many countries—in China at this time, and in the absence of extensive prior negotiation, the project was a non-starter.

Given the constraints of operating in the context of this particular project (distance, lack of trained personnel) it would have been unrealistic to set up a *pre*-initiation baseline review. However, there would have been time, budget, and resources to undertake a *post*-initiation baseline study, in which the UK specialist could have played a facilitating and training role. Such a study would have allowed the implementation team to critique the original plan, and to test its assumptions. They would have quickly come to the conclusion that there were flaws in the project design, and could have fixed things before problems arose. As it was, the project team became very busy going in the wrong direction, and an existing network of municipal teacher training colleges, whose staff the Provincial Education Commission were very keen to have trained, was completely overlooked. Once they had arrived at this impasse, the teams themselves had no data which would help them either to understand the problems they faced, or which they could use in future evaluation.

Conclusions

General conclusions

In this paper I have proposed that there are certain strong advantages to including insiders in project evaluation, and to ensuring that good baseline data is collected. At worst, the chances of instituting safer and more appropriate change processes are higher, while at best, experience to date has been that this approach can help you to get excellent projects, highly motivated teams, and a chance to make a positive and sustainable difference in the operating environment.

Although I have stressed the importance of insider-managed baseline studies, I would wish to avoid any impression of being a born-again baseline zealot! Baseline studies seem important to me because they provide a clear reason for researching the project environment before implementation, are conducive to the inclusion of insiders in their implementation, and seem to lead to better projects. Other kinds of analysis can also contribute to the reduction of risk in social change processes, and each has its place in the project manager's tool-kit, but insider-managed baseline research is cost-effective when compared with outsider-dependent approaches. Also, as we have seen, it brings with it many other advantages.

Extending the reasons for the inclusion of insiders in evaluation

Weir (1995) argued for the inclusion of insiders in evaluation, offering reasons related to:

(a) Outsider objectivity versus insider understanding

(b) Resources

(c) Sample size

(d) Monitoring implementation

(e) Sustainability

I would like to add to these reasons a dimension which connects with Weir's five points—especially his Point 1—but which is not explicitly accounted for in his discussion. I would place this after his first reason for including insiders in project evaluation, and call it 'Ensuring relevance and future evaluation'. As we have seen, baseline evaluation, especially when it includes insider involvement, does not only contribute to the sustainability of a project's benefits: it can also clarify the political and administrative dimensions of a project's environment. In so doing, it makes it possible for project teams to:

- plan safer, more contextually-appropriate projects as a result of *pre-initiation* baseline research
- test and refine existing project plans to ensure a fit between a proposed (outsider-designed) project and its environment
- ensure the existence of an adequate basis for future evaluation.

Insider-informed baseline evaluation, therefore, not only has the advantages that have already been mentioned in association with insider involvement in the whole process of evaluation. It also has the potential

of making sure that educational change processes are appropriate to the contexts in which they are to be carried out.

Received July 1999

A version of this paper was presented at a conference on *National and Regional Issues in ELT: International Perspectives* held in Dhaka, Bangladesh (31 January–2 February 1999).

Notes

1 See, for example, the United Nations Development Programme (UNDP) Evaluation Guidelines (OESP 1997: Chapter 2).

2 For example, the Code of Practice for Grant Funded Services (GFS) ELT Projects currently being developed the British Council, the RUFDATA guidelines for evaluation (M. Saunders and S. Shaw 1998), and the emphasis on impact studies seen in DfID discussions.

3 *Stakeholders* are those who have an interest in a change process. *Primary stakeholders* are those who will be directly affected by the change (e.g. students), *Secondary stakeholders* are those who are involved in implementing or financing the change (e.g. consultants and/or ministry staff), and *External stakeholders* are those who have an interest in the project, but are neither directly affected by the change process nor involved in its implementation.

4 John McGovern and Dianne Wall at IELE, Lancaster University, were the key figures in this process.

5 For most of the Year 12 Project's existence these were Karen Giblin (Latvia), Colin Campbell (Estonia), and Brian McGuire (Lithuania).

6 The Logical Framework is the main planning/reporting instrument used in the Department for International Development Projects. The logical framework uses a four-by-four grid as a concise means for reporting the Aim, Purpose, Outputs, and Activities of a project, showing how the project will be evaluated, and what risks and assumptions have been considered.

References

Alderson, J. C. and **A. Beretta** (eds.). 1992. *Evaluating Second Language Education.* Cambridge: Cambridge University Press.

Alderson, J. C. and **M. Scott.** 1992. 'Insiders, outsiders, and participatory evaluation' in J. C. Alderson and A. Beretta (eds.).

Alderson, J. C. 1992. 'Guidelines for the evaluation of language education' in J. C. Alderson and A. Beretta (eds.).

Bray, T. and **T. Luxon.** 1998. 'The role of baseline studies in ELT projects'. Mimeograph. Lancaster: Lancaster University Institute for English Language Education.

King, J. A., L. L. Morris, and **C. T. Fitz-Gibbon.** 1987. *How to Assess Program Implementation.* Newbury Park, CA: Sage.

Murphy, D., R. Keily, P. Rea-Dickins, and **M. Reid.** 1995. *Guidelines for Evaluation.* London: The British Council.

ODA. 1995. 'Note on enhancing stakeholder participation in aid activities'. London: Department of International Development. www.carryon.oneworld.org/euforic/gb/stake2.htm

OESP. 1997. *Results-Oriented Monitoring and Evaluation: A Handbook for Programme Managers.* New York: OESP Handbook Series, United Nations Development Programme. www.undp.org/undp/eo/monitor.htm

Saunders, M. and **S. Shaw.** 1998. *RUFDATA: A Practical Framework for Evaluation.* London: The British Council.

Thompson, R. M. 1995. *Participatory Project Preparation and Appraisal at the World Bank: The Ugandan Private Sector Competitiveness Project.* Team Technologies. http://www.teamusa.com/part.pdf

Weir, C. and **J. Roberts.** 1994,. Evaluation in ELT. Oxford: Blackwell.

Weir, C. 1995. 'The role of insiders and outsiders in evaluating English language programmes and projects' in T. Crooks and G. Crewes (eds.). *Language and Development.* Jakarta: Indonesia-Australia Language Foundation

The author

Dr Christopher Tribble is a freelance author, lecturer, and consultant. He is an Associate Lecturer at the English Language Centre, King's College, University of London, and an Associate of the Centre for Applied Language Studies, University of Reading. He has published on teaching written communication and the uses of electronic corpora in ELT, and has been a project manager, trainer, and evaluation consultant in Central and Eastern Europe, China, and South Asia for the British Council, the Department for International Development (DfID), Oxfam, and the European Union.

Email: ctribble@lanka.com.uk